Optimising Seismic Array Analysis for Forensic Seismology

Improving the detection and identification of underground nuclear tests using seismic array data

By

NEIL WILKINS

Department of Earth Sciences
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Science.

OCTOBER 2018

Word count: 20,200

Abstract

Seismological techniques have been at the forefront of international efforts to monitor the development and testing of nuclear weapons for more than 70 years. The need for a robust monitoring system has led to the development of seismic instrumentation and data analysis techniques, particularly in arrays of seismometers. Seismic data can be used both to detect clandestine nuclear explosions, and to distinguish them from naturally-occurring earthquakes using discriminants including source depth, regional P/S amplitude ratios and the ratio of body wave magnitude to surface wave magnitude (m_b/M_S).

Seismic arrays have two main functions useful in forensic studies: stacking the seismic traces from each station to improve the signal-to-noise ratio (SNR) of the data, and improving the quality of the signal used for waveform analysis studies. These stacked signals can be applied as vespagrams, where the direction and arrival time of a particular seismic phase can be identified. Stacked signals can also be used in magnitude calculations towards calculating the explosive yield and m_b/M_S for a suspected nuclear test.

In this thesis, the history of forensic seismology is discussed and different stacking techniques applied to seismic data from nuclear tests conducted in North Korea and the results compared for different seismic arrays. I have written a software package VesPy, to perform various useful array analysis functions in Python, which is described in this thesis, alongside example applications.

Finally, these techniques are combined to develop slowness-azimuth station corrections (SASCs) for seismic arrays of the International Monitoring System (IMS) for the region surrounding the North Korean nuclear test site. These corrections are applied to data from five nuclear tests to demonstrate the effect on the signal and magnitude calculations.

I describe a set of recommendations for which stacking methods to use in different situations, and a best practice for developing future SASCs.

Dedication and acknowledgements

I would like to thank everyone who supported me through my PhD project, without whom this thesis would not have been possible. This project was funded by the National Environment Research Council (NERC) and AWE through a NERC CASE scholarship.

Thank you to my supervisor James Wookey for his scientific contributions, helpful advice, and for reading various drafts. Thanks also to Neil Selby, my supervisor at Blacknest, for all his assistance. Thanks to Naomi, Sue, Daniel, Simon and all the staff in Bristol Earth Sciences for helping with all the paperwork, and making life in the department as easy as possible.

I would also like to thank David, Alan, Mike and everyone in the Bristol Geophysics Group at Bristol for being a friendly and helpful group, and for all the biscuits, cakes and even freshly baked bread they provide at the daily coffee breaks. Thanks also to everyone at AWE Blacknest for making me feel at home each time I visited, particularly Ross for giving me a lift each morning!

To office-mates past and present—Jack, Phil, Matt, Antony, Berhe, Khalil, Tom, Tesfaye, Jessica, Rattana—thank you for your help and conversation, and making Room G10b feel like home over the last four years.

To Becky, Ery, Keri, Nicky and Ryan, the best year group, friends and tea-drinking buddies a postgrad student could wish for—I genuinely couldn't have done it without you.

Thanks to my parents, Colin and Christine, my brother Dan, and my nan Rosemary for their support, belief and helping me move several times!

Finally, thanks to Laura, for her love, being there, and keeping me going when it got tough. X

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

Seismology has been used to study nuclear weapons tests for as long as the weapons themselves have existed. During the Cold War, as both sides of the conflict saw the importance of monitoring the other's weapons programme, the development of global seismology occurred in tandem with the development of new nuclear technologies. Nuclear testing has therefore been the catalyst for many seismological techniques during this period, particularly the use of seismic arrays. It also motivated the deployment of global networks of seismometers, from the Worldwide Standardized Seismographic Network (WWSSN) in 1966, to the International Monitoring System (IMS) still being expanded today. Forensic seismology [12], the study of the cause of artificially occurring seismic events, remains at the centre of strategies to monitor the conducting of nuclear tests.

In this chapter, I cover a brief history of the development of forensic seismology and its use in studying nuclear explosions. I explore the background of the Comprehensive Nuclear Test Ban Treaty (CTBT), which aims to stop any further nuclear detonations taking place anywhere on Earth. Finally, I explain the context and background of the specific project I have undertaken as the basis for my PhD.

1.1 Seismology and nuclear testing

The first detonation of a nuclear weapon, "Trinity", was carried out at a purpose-built test site in New Mexico on the 23rd July 1945 as part of the Manhattan Project. The Trinity device, a plutonium-fuelled fission bomb, was placed in a tower 30 m above the New Mexico desert and detonated, exploding with energy equivalent to 20 kilotons (kT) of TNT. American seismologists used portable seismographs to record the ground motion at the test site caused by seismic waves created by the blast [8]. The seismic waves were also recorded at more remote sites: Gutenberg

Table 1.1: Nuclear tests by country

	USA	USSR	UK	France	China	India	Pakistan	DPRK
First test	1945	1949	1952	1960	1964	1974	1998	2006
First underground test	1951	1961	1962	1961	1969	1974	1998	2006
Latest test	1992	1990	1991	1996	1996	1998	1998	2017
No. underground tests	815	496	24	160	24	3	2	6
Total no. tests	1032	715	45	210	47	3	2	6

[37] recorded Pn phase arrivals at seismic observatories in Arizona and California, and used them to measure the origin time of the explosion. Gutenberg's estimate, correct to within two seconds, would be the only accurate recording of the actual detonation time of Trinity, since equipment at the site intended to record the precise shot time had failed.

In 1946, the United States of America conducted the first underwater detonation of an atomic bomb at Bikini Atoll in the Marshall Islands. Seismic waves generated by this explosion were detected by seismographs in California, cementing the idea that seismology could be a useful tool in detecting and studying nuclear explosions from great distances. In 1947, US Army General Dwight Eisenhower directed the United States Air Force to develop the Atomic Energy Detection System (AEDS), tasked with co-ordinating the detection of nuclear explosions anywhere on Earth, with seismology as one of the tools at its disposal. Other important detection techniques for early nuclear tests were acoustic measurements and radiological sampling of the atmosphere and rainwater [13].

In 1952, the US Military deployed seismic instruments around the Nevada Test Site to study the propagation of seismic waves from nuclear explosions [19]. Seven nuclear detonations (called "shots") were conducted both underground and at the surface. The instruments made measurements of the amplitudes and periods of body waves generated by the explosions, and the seismologists made both travel-time and correlation studies of the data.

In 1956, seismologists [9] were able to use seismic data from a British nuclear test at Maralinga, South Australia, to make the first measurements of the thickness of the local continental crust, showing it to be comparable in thickness to the crust in the Canadian and Siberian shield zones. Australia experiences little natural seismicity, so nuclear tests were relatively high-magnitude sources of body waves useful for such studies.

The Partial Test Ban Treaty (PTBT) was ratified by the USA, USSR and UK in 1963, prohibiting nuclear tests from being conducted anywhere but underground (see Section 1.2). Seismology became the most important technique available for detecting these underground explosions. Both the UK Atomic Weapons Research Establishment (AWRE) and the United States government (as part of Project Vela Uniform funding programme) invested in the development of arrays of seismometers for the monitoring of nuclear explosions, and the first seismic arrays were deployed at Eskdalemuir, Scotland and Montana (see Section 1.3). Beyond these array sites, in order to en-

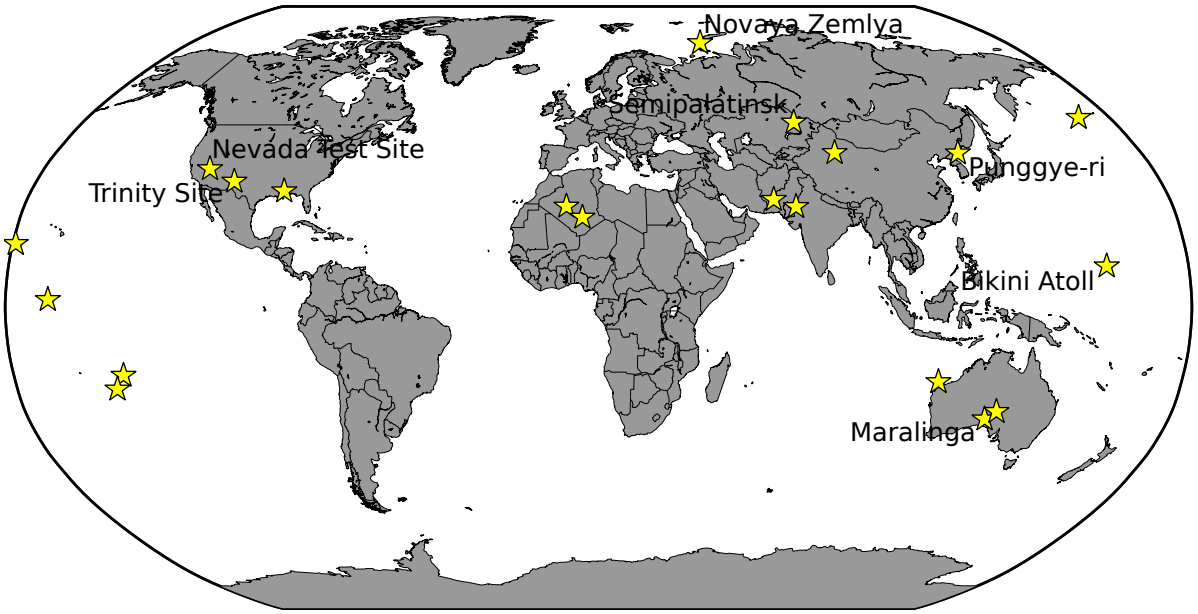

Figure 1.1: Major nuclear testing sites. The Trinity Site in New Mexico is the site of the world's first detonation of an atomic bomb in 1945. A total of 904 tests were conducted by the USA (and UK) at the Nevada Test Site from 1951–1992. In the USSR, 456 explosions were carried out at Semipalatinsk (in modern Kazakhstan) and 130 on the island of Novaya Zemlya (Russia), including a thermonuclear device with an explosive yield of 50 MT, the most powerful nuclear test ever conducted. Punggye-ri was the site of the most recent nuclear tests, conducted by North Korea from 2006–17.

sure a robust monitoring system, highly sensitive seismic instruments would need to be deployed all over the globe. This aligned with the interests of seismologists who had long been interested in setting up a global network of properly calibrated and synchronised seismometers. Following the recommendation from the Advanced Research Projects Agency (ARPA, now DARPA), new standardised instruments were deployed in existing seismic stations, and by 1967 the Worldwide Standardized Seismological Network (WWSSN) comprised 117 stations across the planet, each with three short-period and three long-period seismographs. The network would provide seismic data for geophysical studies of the Earth, not just for nuclear monitoring purposes. As the first global seismic network (in the sense of standardised, synchronised instruments) the WWSSN allowed global seismological studies of a scope and scale that was not possible beforehand. Data from the WWSSN provided accurate earthquake locations and fault plane solutions along plate boundaries that helped identify where these boundaries were [62], and which way plates were moving relative to each other [58], developing the theory of Plate Tectonics.

In the area of nuclear test monitoring, progress towards a Comprehensive Test Ban Treaty (CTBT) depended on the ability to identify which seismic events were explosions. Early methods for distinguishing between an underground explosion and an earthquake depended on the "first motion" criterion [27], that is, the direction of the first motion of the seismograph trace for the

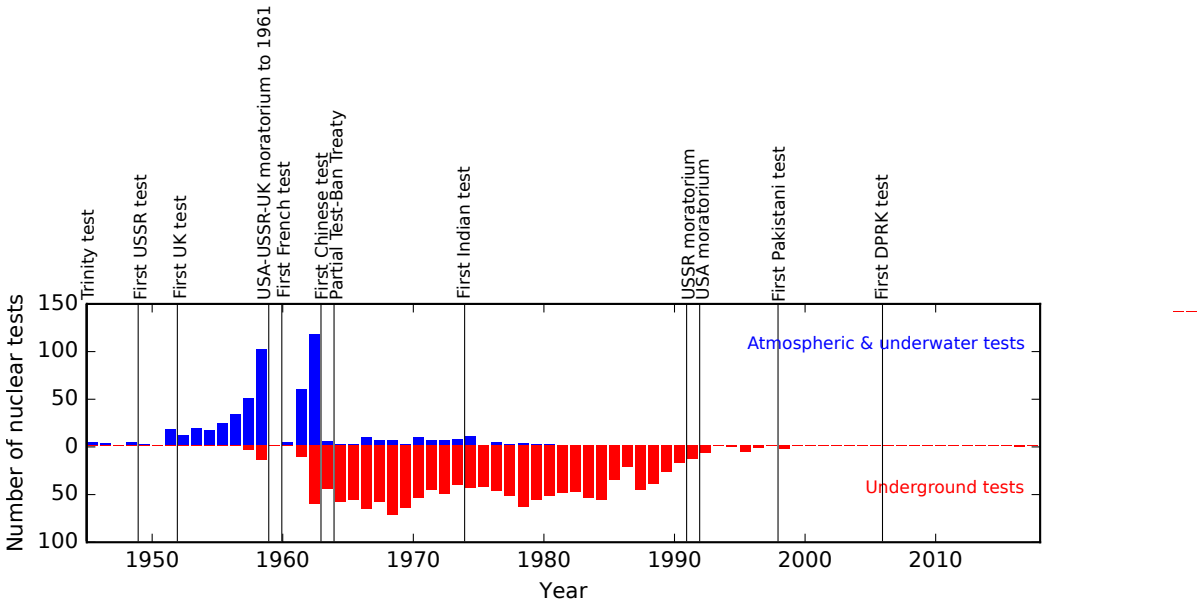

Figure 1.2: Nuclear tests conducted by year. The nuclear age began in July 1945 when the USA detonated the Trinity device in New Mexico. A total of 2058 known nuclear devices have been detonated between 1945 and September 2017. A temporary moratorium was observed by the USA, USSR and UK from November 1958 to September 1961. In 1963, the Partial Test-Ban Treaty was signed by the USA, USSR and UK, prohibiting those countries from making any further atmospheric or underwater tests. The USSR/Russian Federation and the USA have observed moratoria against any nuclear tests since January 1991 and October 1992, respectively.

P-wave arrival from the event. An explosion was assumed to be a purely compressive source, leading to the first motion always being upwards. Earthquakes, on the other hand, have a source with motion along a fault plane, and so the first motion could be up or down depending on the direction to the source. The first motion on a seismic trace is often difficult to determine and subject to interpretation by the analyst looking at the data. Another way of ruling out an explosion (a *negative* discriminant) is the depth of the source, which could be determined using the P-wave and pP, its reflection at the Earth's surface. Detection of pP can be obscured by multiple reflections caused by layering at the surface [8]. The first *positive* discriminant that could identify an explosion was the m_b/M_S criterion, first suggested in 1963 and well-established by the late 1960s [50]. Research at AWRE Blacknest indicated that explosions were found to produce Rayleigh waves with amplitudes of around an order of magnitude less than earthquakes of the same body wave magnitude [8].

However, attempts to use seismology to verify compliance with the 1974 Threshold Test Ban Treaty (TTBT) between the USA and USSR faced two main obstacles. The first was the difficulty of accurately estimating yield from seismic data. The TTBT set the maximum permitted nuclear warhead with an explosive force equivalent to 150 kT of TNT. Studies of American nuclear tests (of which the explosive yield was known) had produced a relationship between the amplitude

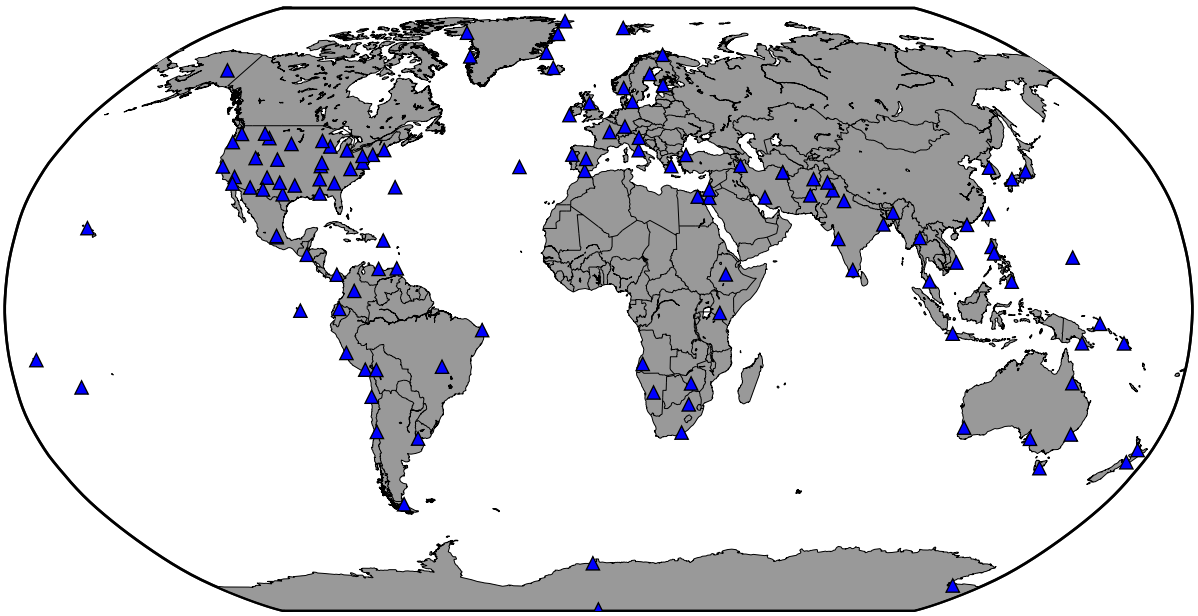

Figure 1.3: Stations of the Worldwide Standardized Seismographic Network. Each station had three short-period and three long-period seismometers to record the north-south, east-west and vertical components of ground motion.

recorded in the seismogram and the energy of the explosion. A complicating issue was that underground nuclear explosions could be decoupled from the Earth by detonating them within a large pre-excavated cavity, reducing the amplitude of the seismic waves generated by the blast [78]. The second issue was the magnitude threshold below which a seismic event could be identified as an explosion. In 1968 the Stockholm International Peace Research Institute (SIPRI) produced a report on the state of international monitoring capabilities. The report found that both Soviet networks and the WWSSN were able to locate and identify explosions (using the m_b/M_S criterion) with yields as low as 20–40 kT in hard rock. However disagreement over the magnitude below which identification could be made with enough certainty continued through the 1970s, frustrating efforts to negotiate a CTBT [8].

Nevertheless, seismologists were able to use seismic waves from underground nuclear tests, as well as the instrumentation now available, to make many scientific advances in the study of Earth structure. An underground explosion can present a useful seismic source for a global seismologist in that the precise location and origin time can be known, and provide strong P-wave signals for use in travel-time studies. The large number of tests conducted at the Nevada Test Site produced a large quantity of seismic data useful for studying the crustal structure of Nevada and Arizona [26], and the lithospheric structure of California. P'P' reflections from nuclear tests conducted at Novaya Zemlya in 1970–1 gave an early glimpse of the structure of the upper mantle beneath Antarctica [1]. Ground motion from seismographs deployed in the Las Vegas Valley provided useful data for structural engineers [53].

Array design (see Section 1.3) and political negotiations towards a CTBT (Section 1.2) continued throughout the 1970s and 1980s, with the final treaty opening for signatures in 1996. The establishment of the Preparatory Commission for the CTBT Organisation (CTBTO) and the International Data Centre (IDC) in Vienna co-ordinated international monitoring efforts for the first time in one central organisation. The CTBTO have created the International Monitoring System (IMS), which includes not just seismic but also infrasound, hydroacoustic and radionuclide monitoring stations to detect nuclear weapons tests occurring anywhere on the planet. While the treaty itself is not yet in effect, data from the IMS has been useful in monitoring the six announced nuclear weapons tests in North Korea, the only tests conducted since the CTBT was finalised. IMS data has also been used to study the Chelyabinsk meteorite impact in Russia in 2013 [28], the dispersal of radioactive material from the Fukushima Daiichi nuclear power plant disaster in 2011 [52] and to aid in the search for the missing Argentine submarine ARA San Juan in 2017 [21].

Forensic seismology remains an important technique for monitoring nuclear testing, alongside infrasound, hydroacoustic, radionuclide, synthetic aperture radar and satellite imagery.

1.1.1 Seismic discriminants for underground explosions

There are three main seismic discriminants used in identifying underground nuclear explosions.

1. **Depth:** if the depth of a seismic event can be established as being greater than 10 km, it can be discounted as being of manmade origin due to the impracticability of emplacing a nuclear device at this depth. Underground nuclear tests generally have an overburden on the order of hundreds of metres (or around 400 ft times the cube root of the yield [80]) to contain any leakage of radioactive material from the explosion. The depth of focus of an event can be determined from crustal phases [25], or the spectral properties of the waveform [39].

2. **Regional amplitude ratio of P- to S-waves:** explosions are observed to generate weaker S-waves compared to P-waves [79]; for most natural earthquakes we expect the inverse. The amplitude of body waves around 2 Hz is measured at regional (<2000 km) of the source to identify possible explosions.

3. m_b/M_S **Magnitude ratio:** underground explosions produce relatively weaker surface waves compared with natural earthquakes ([50],[76]). The ratio of m_b, measured for the first P-wave arrival around 1 Hz to M_S, measured for L_g waves can be used to discriminate between explosions and earthquakes. The IDC uses a screening line of $M_S = 1.25m_b - 2.2$ [30], with events above this line being identified as earthquakes, and those below as explosions. However, for the DPRK nuclear tests, measurements of m_b/M_S have been on the threshold, leading to uncertainty in the classification of the events from the seismic

magnitude ratio. I reproduce the figure of Bowers & Selby (2009) [12], showing m_b vs. M_S, and the screening line for a number of earthquakes.

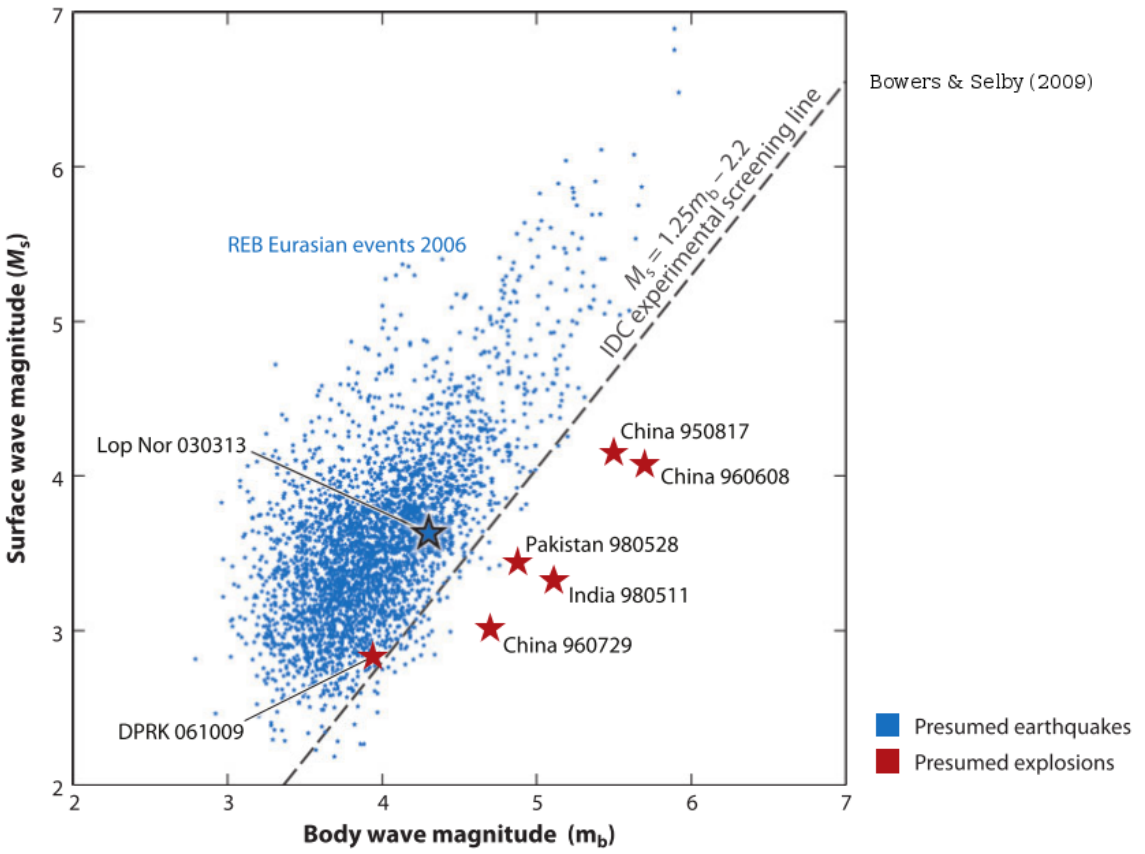

Figure 1.4: From Bowers & Selby (2009) [12], a graph of m_b vs. M_S for underground explosions and earthquakes in Eurasia, including the announced DPRK test of 2006, announced nuclear tests in India, Pakistan and China, and the earthquake of 2003 in Lop Nor, China, which was suspected of being a nuclear explosion. The 2006 DPRK event sits on the screening line, making the m_b/M_S criterion ambiguous for this event.

Broadband seismic data from regional stations can use a combination of these discriminants to identify suspicious events in real time [63].

1.2 Nuclear test-ban treaties

Since 1945, the United Nations has attempted to regulate the proliferation of nuclear armaments, establishing the Atomic Energy Commission (UNAEC) to address problems associated with the new technology. Over the next several decades of the Cold War, proposals were made from both sides towards nuclear disarmament and the control of fissile material, but no agreement was

reached for over two decades. Grassroots peace movements, such as the Campaign for Nuclear Disarmament (CND) came to prominence in the 1950s and 1960s as the public reacted to the perceived increase in the risk of nuclear war between the United States and the Soviet Union.

As well as these global security concerns, the 1950s saw an increase in the scientific awareness of the harmful effects of radioactive material released into the biosphere by nuclear testing. In 1954 the crew of the Japanese fishing vessel *Daigo Fukuryuu Maru* (Lucky Dragon No. 5) were exposed to radioactive debris from US nuclear testing in the Marshall Islands. All members of the crew suffered from Acute Radiation Syndrome (ARS), eventually recovering after several weeks, with the exception of the radioman who died from complications arising from his treatment. The thermonuclear device tested had produced much more fallout than had been predicted, leading to increased scrutiny on the environmental effects of atmospheric tests.

The Indian Prime Minister Jawaharlal Nehru called for a "standstill agreement" on nuclear testing in 1954. However it was not until 1963, a year after the heightened tensions surrounding the Cuban Missile Crisis, that the first agreement, the Partial Test Ban Treaty (1963) would be established. The PTBT prohibits all nuclear detonations in the atmosphere, underwater or in outer space, meaning all future testing would have to be conducted underground. The treaty was rapidly ratified by the USSR, USA and UK and was made effective in October 1963. However the new nuclear powers of France and China, having only tested their first weapons within the previous three years, did not sign the treaty. While the PTBT failed to prevent the proliferation and continued development of nuclear weapons, the shift to underground testing significantly reduced the amount of radioactive particles released into the atmosphere.

The PTBT was followed in 1968 by the Treaty on the Non-Proliferation of Nuclear Weapons (NPT), which sought to prevent the spread of weapons technology beyond the existing nuclear powers. The NPT prevented its signatories from transferring nuclear weapons to other countries, or from pursuing their own weapons programmes. Instead, it encouraged the spread of peaceful nuclear technologies, such as fission power. However, critics of the treaty observed that the existing nuclear powers did little to disarm their own nuclear arsenals, as the treaty required. Non-signatories of the NPT include India, Israel and Pakistan.

The USA and USSR would continue discussions to limit the development of more powerful weapons, resulting in the bilateral Threshold Test Ban Treaty (TTBT) in 1974 being signed between the two countries. Under this agreement, neither country would test a nuclear device with a yield exceeding 150 kT. Importantly, the TTBT required the two states to exchange technical data to assist the verification of the treaty. However the accurate estimation of the explosive yield of an underground nuclear test still presented many technical problems that neither side was confident had been overcome. 1974 also saw the first test of a nuclear weapon by India. To improve confidence in verification methods, the USA and USSR co-operated in the Joint Verification Experiment in 1988 to evaluate yield measurement systems at both the Nevada Test Site and the Soviet test site in Semipalatinsk, Kazakhstan [6].

Discussions would continue into the 1990s to amend the international PTBT and work towards a Comprehensive Test Ban Treaty (CTBT) that would prohibit all future nuclear tests worldwide. The USSR declared a moratorium on nuclear tests in 1991, and the USA and UK had followed suit by the end of 1992. The terms of the CTBT were finalised in 1996 and the Treaty opened for signatures. In addition, the Preparatory Commission for the CTBT Organisation (CTBTO) was established in Vienna to co-ordinate the verification of the treaty when it enters into force.

The CTBT almost immediately faced obstacles to its entry into force, with renewed nuclear testing in India on 13th May 1998, quickly followed by the first nuclear tests in Pakistan on the 29th of the same month. In 1999, the US Senate rejected the treaty, and despite it having been signed by President Bill Clinton in 1996, the USA has still yet to ratify the CTBT. More recently, the Democratic People's Republic of Korea (DPRK) has tested a series of nuclear devices between 2006 and 2017, leading to increased political tension in East Asia and renewed public fears of nuclear war. In April 2018, North and South Korea signed the Panmunjom Declaration, which included confirmation of "the common goal of realizing, through complete denuclearization, a nuclear-free Korean Peninsula". This statement was reaffirmed in a joint USA-DPRK statement in June 2018 with a commitment "to work towards the complete denuclearization of the Korean Peninsula". The DPRK has since declared their nuclear test site at Punggye-ri shut down, and has conducted controlled demolitions of tunnels at the site [59].

At the time of writing, the CTBT is still far from entry into force. Of the nations required to ratify the treaty before it is in effect, three have not signed at all: the DPRK, India and Pakistan. In addition, five key signatories have yet to fully ratify the treaty: China, Egypt, Iran, Israel and the USA. However, the CTBTO continues to push for its adoption, and operates the International Monitoring System to monitor the planet for nuclear testing.

In 2017, the United Nations began negotiations for the first international agreement to prohibit all nuclear weapons, the Nuclear Weapon Ban Treaty (NWBT). Opened for signatures in September 2017, the treaty aims to prohibit not just the testing and production of nuclear weapons, but their stockpiling and any use at all, requiring its signatories to work towards the complete elimination of their nuclear arsenal. No nuclear weapons state has signed the NWBT to date.

1.3 Seismic arrays

In 1959 the British Atomic Weapons Research Establishment (AWRE) assembled a group of scientists with a background in radar to begin experimenting with techniques to improve seismological observations of nuclear tests. The operation of a seismic array is based on the same principles a radio antenna uses to enhance radio signals with respect to background noise, and use the propagation of the wavefronts across the array to locate the source of the signal. The first

Table 1.2: Nuclear test-ban and non-proliferation treaties

Year signed	Effective	Treaty	Provision
1963	1963	Partial Test Ban Treaty	No nuclear tests conducted except underground
1968	1970	Non-Proliferation Treaty	Obligation towards proliferation and disarmament, but the right to use peaceful nuclear technology
1974	1990	Threshold Test Ban Treaty	No nuclear tests conducted with yield above 150 kT
1996	Not in effect	Comprehensive Test Ban Treaty	No nuclear tests conducted
2017	Not in effect	Nuclear Weapon Ban Treaty	Elimination of nuclear weapons

small seismic array was deployed on Salisbury Plain in 1960–1. The array design was a success, but the location in southern England was high in background noise. It was decided that a linear array of seismometers would be established in a remote area of Scotland with suitable geology. This was the Eskdalemuir seismic array (EKA) [77]. This array comprises two perpendicular arms of around 10 km length. Despite still being a relatively noisy site, the array was found to improve the seismic signal-to-noise ratio (SNR) [82]. This successful cross-shaped design, called the UKAEA geometry, was also used in arrays installed in Yellowknife (Canada), Warramunga (Australia), Gauribidanur (India) and Brasília (Brazil).

The success of the British seismometer arrays inspired the United States to deploy the 200 km Large-Aperture Seismic Array (LASA), the first of its size, in Montana. Funded by the Vela program, in which the US government provided millions of dollars towards improving seismological infrastructure for the monitoring of nuclear tests, LASA consisted of 625 instruments arranged in 21 sub-arrays deployed in a series of concentric rings [36]. Although the instruments were installed by the end of 1964, the array would not be fully operational until 1971 when IBM's data-processing software was complete. Similar designs, and additional US funding, established the NORSAR array in Norway, and the ALPA array in Alaska.

Data from the newly-established UK and US arrays proved useful to seismologists studying the structure of the Earth. Seismic data recorded at LASA from a nuclear test conducted at the Nevada Test Site was used to detect PKIIKP phases and make a measurement of the P-wave velocity in the inner core [9].

In the late 1960s, the Advanced Research Projects Agency (ARPA) established ARPANET, the first computer network to use the TCP/IP transfer protocol to exchange data between US research institutes across the country. In 1969, NORSAR became the first site outside the USA to be connected to ARPANET for the transmission of seismic data. The technology behind ARPANET would be used to establish the infrastructure of the internet in the early 1990s, and the ARPANET connection at NORSAR was the forerunner to the instant global transfer of data from seismic

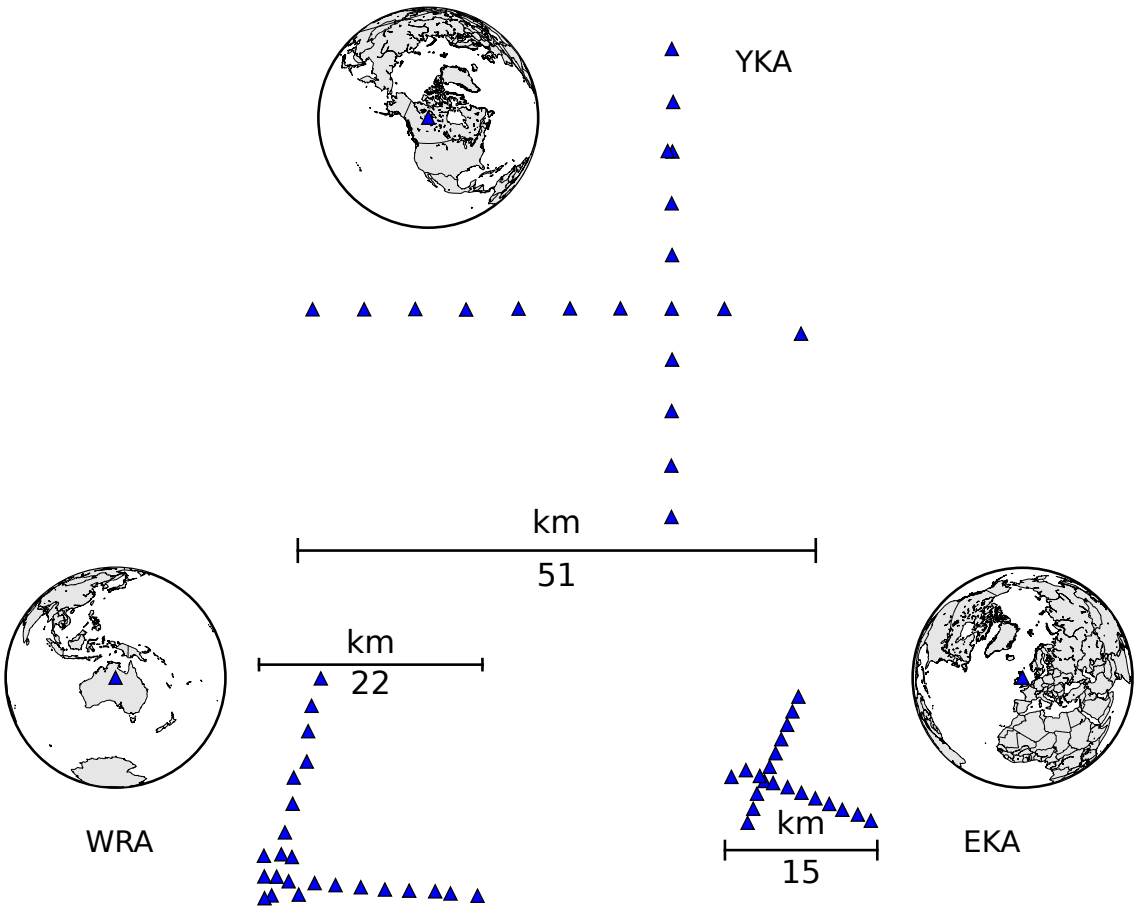

Figure 1.5: UKAEA type arrays. Clockwise from top: YKA, deployed near Yellowknife, Canada; EKA, near Eskdalemuir, Scotland; and WRA, near Warramunga, Australia. In the 1960s the arrays consisted of analogue short-period instruments, which have since been replaced with digital seismometers. Eskdalemuir remains the longest-continuously-operational steerable seismic array in the world.

networks around the world that underpins much of modern seismology [55].

The next advance in seismic array design came with the development of borehole instruments during the 1970s. Seismic arrays deployed at depths of 50–100 m are subject to significantly less noise than instruments placed at the surface. During the same period the development of ocean-bottom seismometers (OBS) allowed seismological data to be collected from the oceans for the first time.

The mid-1970s saw the arrival of digital instruments in seismology. One key advantage of digital seismometers is that they allow for a much greater dynamic range in the signals they can record. Earlier analogue instruments, particularly those used in the West, were restricted

to short- or long-period passbands to ensure that signals were not drowned out by higher-amplitude noise. Broadband digital seismometers are able to record a much wider spectrum of the incoming wavefield, and still resolve small amplitude signals within their frequency band. The Gräfenberg array, deployed in southern Germany in 1976, was the first continuously-recording digital broadband seismometer array.

Despite these technological advances, it was realised that a few very large arrays like LASA and NORSAR would be insufficient for test-ban monitoring purposes. The operation of these arrays also proved to be prohibitively expensive and in 1978 LASA was shut down, and NORSAR reduced in size. The 1980s saw an increased interest in smaller-aperture seismic arrays, such as NORESS and ARCESS in Norway, equipped with digital instruments that had become more economical.

Primary seismic arrays deployed as part of the International Monitoring System (IMS) have at least nine vertical short-period insturments, and at least one three-component broadband sensor [18]. Stations are deployed in vaults or 50 m boreholes, in sites chosen to be in geological homogeneous areas with low cultural noise, such as the TORD array deployed in Niger in 2005. Primary seismic stations transmit continuous data to the International Data Centre (IDC) in real time.

Modern seismometers, with digital sensors continuously connected to the internet allow almost instant collection of data, and the opportunity to investigate suspicious events with unprecedented speed. The announced underground nuclear test conducted by the DPRK in September 2017 was preliminarily verified within minutes by observatories in Japan and China.

The 21st century has also seen the deployment of seismic arrays on the scale of entire countries. The High Sensitivity Seismic Network (HI-NET) consists of around 600 seismic stations deployed across Japan from 2000-2005, for the purpose of monitoring seismic activity throughout the country. The United States Transportable Array (USArray) consists of temporary seismic stations rolled out across the contiguous United States, followed by Alaska. Each station remained in place for one year before being moved on, with the array at any point forming a wide north-south strip of stations with a spacing of 50 km.

1.3.1 Array techniques

The establishment of seismic arrays enabled the development of new data processing and analysis techniques.

The basic theory of beamforming and summing time series data from each station to improve signal-to-noise ratio were well established by radio engineers working with radar arrays during the 1940s. These techniques were adapted to seismic data, by e.g. Robinson [65].

With the construction of the Large Aperture Seismic Array (LASA), seismologists were able to apply array analysis techniques on seismic data. In order to improve the signal-to-noise ratio (SNR) of the seismic data, beamforming techniques including delay-and-sum, weighted sum

and filtered sum were used [15]. The delay-and-sum method (see Section 2.1) simply sums the amplitude of the seismogram recorded at each station in the array after accounting for the delay in arrival as the waves propagate across the array. The weighted sum and filtered sum methods apply an amplitude weighting and filtering function to the contribution from each sub-array based on background noise studies carried out in the vicinity. Other signal processing techniques, such as cross-covariance and power spectral analysis were also applied at LASA [3].

Beamforming in the frequency domain (FK analysis) was developed by Capon in the late 1960s and illustrated using data from LASA [14]. This technique is described in more detail in Chapter 3. The ability of a seismic array to accurately detect and locate the origin of seismic waves depends on the array's geometry, which controls the response of the array in the frequency domain. This response can be modelled by the array response function (ARF), given by

$$\left|A(\vec{k})\right|^2 = \frac{1}{N}\sum_{j=1}^{N} e^{i\vec{k}\vec{r}_j}, \tag{1.1}$$

where $\vec{k}$ is the vector describing the wavenumber of the incoming seismic waves and $\vec{r}_j$ is the vector describing the position of the jth instrument in the array from some reference point [14]. The ARF gives the power recorded by the array for incoming seismic waves travelling with a particular slowness vector. A large aperture array, for instance, has a better resolution for distinguishing between seismic phases arriving with different slownesses. An array with instruments deployed in concentric rings gives a more even response across seismic waves arriving from different directions (backazimuths).

UKAEA arrays such as ESK and YKA, with their two-sided geometries have ARFs with strong side lobes. This means that events producing seismic waves which arrive at the array away from these directions are detected with poorer resolution. To overcome this, researchers at AWRE Blacknest [5] developed a correlation method whereby beams calculated for each arm of the array were cross-correlated to detect coherent signals that were in phase across both arms.

Velocity Spectral Analysis (VESPA) was developed by Davies *et al.* [22], using array beams calculated at LASA for a range of different slownesses to distinguish two seismic phases arriving from different directions at different velocities simultaneously. VESPA is one of the techniques implemented in the VesPy software package, and is described more fully in Chapter 3.

Other signal stacking techniques and detection statistics developed for seismic array analysis are described in Chapter 2.

1.4 Aims and Objectives

The objective of this PhD thesis is to develop methods and software for seismic array analysis, to be applied specifically to the problem of forensic seismology, as well as to other applications.

In Chapter 2, I describe in more detail the various time-domain stacking techniques that can be applied at seismic arrays. I then compare the results of different methods for a variety of datasets with different properties.

In Chapter 3 I describe VesPy, the Python software package I have developed during my PhD studies, to provide seismologists with a set of useful functions they can apply to seismic array data for a variety of scales and applications.

Chapter 4 goes into more detail about the nuclear tests conducted in the DPRK in the last 12 years. I apply the methods incorporated in VesPy to develop a series of corrections to the predicted slowness of arrivals at IMS seismic arrays in the East Asia region. These corrections improve the signal enhancement of the time-series stack, and can be shown to decrease the uncertainty in magnitude estimates obtained from the stacked data. These corrections can thus have a beneficial impact on the identification of nuclear testing in the future.

Finally, Chapter 5 contains concluding remarks about the findings of my project. I make suggestions for future forensic studies: which seismic arrays and which stacking techniques are applicable in each situation, and a best practice for developing future slowness corrections for seismic arrays. I also suggest some further work that could extend these studies beyond this project.

CHAPTER 2

A COMPARISON OF STACKING METHODOLOGIES FOR FORENSIC SEISMOLOGY

Seismic arrays were first deployed in the 1960s as high-quality seismic data was sought after, particularly for application in detecting and monitoring nuclear weapons testing on both sides of the Cold War [65]. An array is defined as a number of networked seismic instruments deployed such that they each report continuous data, with synchronised time stamps, to the same operator [70]. Two-dimensional deployments of arrays allow for beamforming techniques, where the signals from each array element are stacked to improve the quality of the signal. The stacked signal, which has a higher signal-to-noise ratio (SNR) than the individual traces, is called the "beam". The term beamforming can thus refer to the process of computing the stacked signals themselves, or to methods where directional information about the source of the signal can be inferred (as with vespagrams and fk analysis). While signal-enhancing and beamforming methods were the original purpose behind the early development of arrays of analogue seismometers [5], digital array data and modern computing power have made more advanced signal-processing methods possible, such as cross-correlation beamforming [67] and detection [34]. Signal stacking of time series data will be the focus of this chapter.

The principles of stacking can be applied on a range of scales of seismic study. Localised, dense arrays of seismometers or geophones provide high resolution data at oil exploration sites, and stacking is used to detect low magnitude seismic events that occur in the reservoir [73]. At the other end of the scale, stacked data from large aperture teleseismic arrays are used to detect and identify events with sources many thousands of kilometres away. Stacked data from a global distribution of seismometers has been used to estimate the explosive yield of the September 2017 announced nuclear test in North Korea from the body wave magnitude derived from the teleseismic P-wave [16].

In this chapter, I describe a number of methods used to stack time series data. I then focus on the specific application of these techniques to forensic seismology. I apply a number of stacking methods to seismic data recorded at regional and teleseismic distances at arrays of different size and with different signal coherence properties to explore the most effective technique in each situation. I contrast methods useful in seismic waveform studies and in the detection of seismic events, and in the final section, make recommendations for which techniques to use in different circumstances.

2.1 Stacking Techniques

The stacking of time series data from seismic arrays is a common technique used to enhance the signal-to-noise ratio (SNR) of seismic data. With modern digital seismic records and synchronised timestamps across each instrument within the array, stacking is computationally quick to implement while potentially achieving a significant SNR increase.

While many stacking procedures of various computational complexities exist, all are based on the fundamental assumption that, within the relevant timespan, only one coherent signal is recorded as it propagates across the array, and the rest of the seismogram contains only incoherent noise. Where this approximation fails, and coherent noise of significant amplitude contaminates the seismic data, the SNR improvement achieved by stacking is reduced.

Another key assumption is that the wavefronts propagate across the array as parallel linear wavefronts. Under this plane wave approximation, the seismic phases arrive at each station in the array with the same slowness. The time shifts for each station depends only on the position of each station with respect to a fixed reference point. An exception is the adaptive stack (Section 2.1.2), which takes the planar wavefronts of the linear stack as its starting point, but seeks to iteratively improve on the time shifts over simply using a velocity model.

2.1.1 Linear stack

The simplest method for stacking is the linear delay and sum stack, which shifts each trace to account for the delayed arrival of a particular phase as it propagates across the array, and then takes the average amplitude of the shifted traces. The delay time τ_j for a seismic phase propagating with horizontal slowness $\vec{s}$ at station j located at a displacement $\vec{r}_j$ from the array reference point, is given by

$$\tau_j = \vec{r}_j \cdot \vec{s}. \tag{2.1}$$

The linear stack v for an array of N seismometers is given by

$$v(t) = \frac{1}{N} \sum_{j=1}^{N} u_j(t - \tau_j), \tag{2.2}$$

where $u_j(t)$ are the shifted seismic traces, i.e. the amplitude recorded at station j at time t, correcting for the moveout of the wavefield across the array. In the ideal case where the noise is incoherent and random, the signal to noise ratio of the stack is improved by $\sqrt{N}$ with respect to the unstacked traces. An additional advantage of the linear stack is that it preserves the shape of the signal waveform, allowing the beam, with its improved SNR, to be used for waveform analysis [66].

2.1.2 Adaptive stack

The adaptive stack method of Rawlinson and Kennett [64] uses an iterative process to minimise the misfit between the stack and the individual traces. Instead of relying on an *a priori* velocity model to provide delay times with which to shift the traces, the traces are shifted in time and scaled in amplitude so that they minimise the L_p misfit, which is done using a direct search algorithm over the additional shift α.

The misfit is defined as

$$\mathscr{L}_p = \sum_{k=1}^{T} \left| v(t_k) - \alpha_j u_j(t_k - \tau_j - \beta_j). \right|^p \tag{2.3}$$

The linear stack $v(t)$ is then recalculated with the adjusted amplitude scalings α_j and time shifts $(\tau_j + \beta_j)$, and the process is repeated until the alignment of the traces is stable over repeated iterations.

The resulting stacked beam is the one which produces the smallest misfit, and provided the waveforms at different stations have similar shapes, the traces are aligned beyond the 1D velocity model, accounting for small perturbations in the local velocity structure to the array. These revised shifts and scalings are often called static corrections.

2.1.3 nth-root stack

Non-linear stacking techniques are able to improve the SNR beyond the $\sqrt{N}$ limit. One such technique is the nth-root stack of Kanasewich *et al.* [43], so called because it involves taking the nth root of the amplitude of the data before summing, where $n > 1$ is an integer. This reduces the difference in amplitude at different points in the time series, diminishing the importance of periods of larger amplitude in the amplitude of the stack. Instead the coherence of the signal across the array is more important.

In the first step, the nth root is taken, preserving sign, and the traces are summed as follows

$$v'_n(t) = \frac{1}{N} \sum_{j=1}^{N} u_j(t - \tau_j)^{\frac{1}{n}} \operatorname{sign}\left[u_j(t - \tau_j)\right] \tag{2.4}$$

The resulting stack is then raised back to the nth power to produced the nth-root stack.

$$v_n(t) = \left|v'_n(t)\right|^n \operatorname{sign}\left[u_j(t-\tau_j)\right] \tag{2.5}$$

The nth-root stack, as a non-linear process, does not preserve the shape of the waveform beyond its polarity, so it is not suitable for many studies of the waveform itself. These include shear-wave splitting [74] or receiver function studies [68], and in forensic seismology, calculations of the P/S amplitude ratio as an event discriminant [60].

2.1.4 Phase-weighted stack

The phase-weighted stack, first formulated by Schimmel and Paulssen [69], makes a direct estimate of the instantaneous phase of the traces at each point in time and uses this to weight the contribution of that trace to the final stack. It thus explicitly uses the coherence of the signal, rather than the randomness of the noise, to improve the SNR.

The phase estimate is given by

$$\phi_j(t) = \arg[u_j(t) + iH(u_j)(t)], \tag{2.6}$$

where $H(u_j)(t)$ is the Hilbert transform of the seismic traces u_j.

The phase-weighted stack is then calculated as follows, with ν an integer controlling the sharpness of the phase-weighting:

$$v_\phi(t) = \frac{1}{N}\sum_{j=1}^{N} u_j(t-\tau_j)\left|\frac{1}{N}\sum_{j=1}^{N} e^{i\phi_j(t-\tau_j)}\right|^\nu. \tag{2.7}$$

2.1.5 Statistical stacks

It is also possible to weight the stack using a different statistic, such as the semblance or the F-statistic, as a measure of the coherence.

The semblance $S(t)$ estimates the normalised power in the linear stack as

$$S(t) = \frac{\left(\sum_{j=1}^{N} u_j(t)\right)^2}{N\sum_{j=1}^{N} u_j^2(t)}. \tag{2.8}$$

The semblance can then be used as a weighting function applied to the linear stack:

$$v_S(t) = S(t)v(t). \tag{2.9}$$

The semblance (or a similar measurement of coherence) has been used across three-component seismic data to use the entire seismic wavefield in the stacking process [44].

Related to the semblance, the F-statistic, formulated by Blandford [7], provides a measure of the instantaneous signal-to-noise ratio of the traces, and so can also be used as a weighting function. It is given by

$$F(t) = \frac{N(N-1)v(t)^2}{\sum_{j=1}^{N} u_j(t-\tau_j)^2 - Nv(t)^2}, \tag{2.10}$$

F can be reformulated in terms of the semblance as follows [11]:

$$F(t) = (N-1)\frac{S(t)}{1-S(t)}. \tag{2.11}$$

The form of F in Equation 3.8 requires that the noise be incoherent between the different array elements. Selby [72] formulated a generalised F for arrays with coherent noise.

2.2 Implementation

All the above stacking methods have been implemented in the Python programming language, with most of them now released as part of the VesPy software package (described in full in Chapter 3). The functions use ObsPy [4] seismic data handling structures for their input, and NumPy's numerical arrays [81] for their output. For the adaptive stack, I used the `sactcas` module, written in the Fortran 77 programming language, and adapted by Helffrich and Wookey from the original `tcas` code of Rawlinson *et al.* [64] to be made compatible with the SAC data format.

2.3 The order of nth-root stacking

The choice of n, the order of the nth-root stack is important in controlling the shape of the resulting stacked waveform. $n = 1$ simply gives the linear stack. As n increases, the effect of the suppression of amplitude differences throughout the signal gives a greater weight to the first arrivals of the seismic phases, leading to an increased "spikiness" in the waveform. The dramatically increased amplitude of the signal with respect to the background noise makes higher-order nth-root stacks more suitable as event detectors, and they work well in vespagrams (see Section 3.2.4) for capturing phase arrivals at seismic arrays.

To demonstrate the effect of changing n, I apply the nth-root stack process in vespagrams for the 2006 announced nuclear test in the DPRK. As the lowest-magnitude announced nuclear test in North Korea, the 2006 event represents a situation in which signal-to-noise improvement is more important for detection purposes than for the later, larger magnitude events. I generate vespagrams for a regional array, MJAR (at a distance of 8° from the event), for a teleseismic array, WRA (61°) and at a teleseismic very large array, a subset of 48 stations of the US Transportable

Array (USArray), in the northeastern USA (~ 80°). The arrays and their locations relative to the site of the DPRK events are shown in Figures 2.1–2.3.

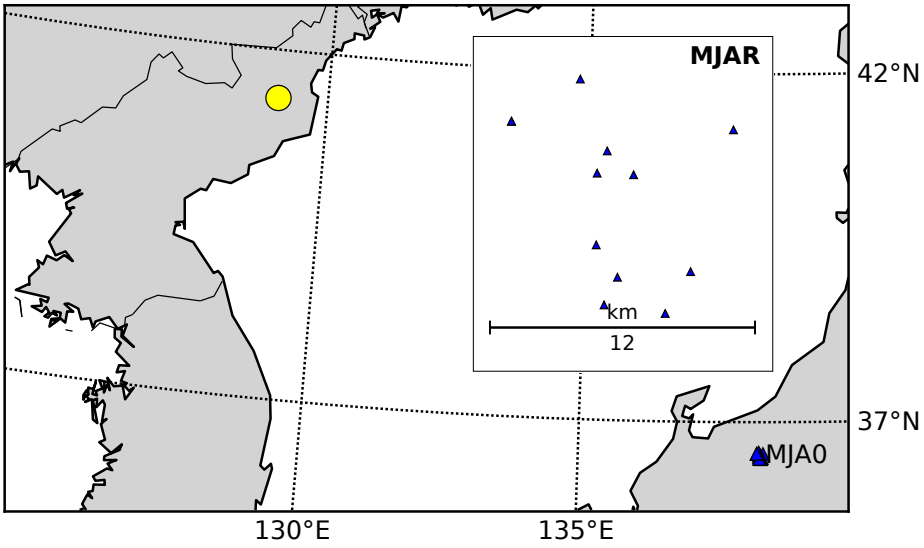

Figure 2.1: Location of the DPRK announced nuclear tests and the Matsushiro Seismic Array (MJAR) in Japan. Inset: array geometry of MJAR.

For each array, the waveforms are filtered between 1.0 and 5.0 Hz to sample the body waves. The power in the nth-root stack with $n = 1, 4, 8$ is calculated in a moving time window of 0.5 s on a segment of the waveform of 20 s around the predicted arrival time. The delay times are calculated using the *ak135* model [46]. The vespagrams plot this power as a function of time and backazimuth, with regions of high power corresponding to the backazimuth and arrival time of incident seismic energy. Figures 2.4–2.6 show the results for all three arrays.

At each array, increasing the order of the nth-root stack improves the noise suppression in the vespagram, which in the case of MJAR allows the P-wave to be identified in the fourth-order vespagram where it cannot be seen in the first-order (Figure 2.4). In addition, the spread of the power maximum in the vertical direction is reduced in the higher-order vespagrams. Increasing n thus improves the backazimuth resolution of the vespagram, allowing the backazimuth to be identified with greater precision. However, in the case of MJAR, a later arrival in the eighth-order vespagram causes the P-wave to be obscured. This problem may be mitigated if the time window of a suspected phase arrival is narrowed, but raises a problem for the detection of seismic phases over a longer period of time.

Such vespagrams can be produced to monitor seismic arrivals from a given site. For a particular region of interest, the vespagram can be calculated for a backazimuth range around that of the suspected test site, using a theoretical value for the slowness of the P-wave arrival. Similarly, a slowness vespagram could be produced with a fixed backazimuth for the region.

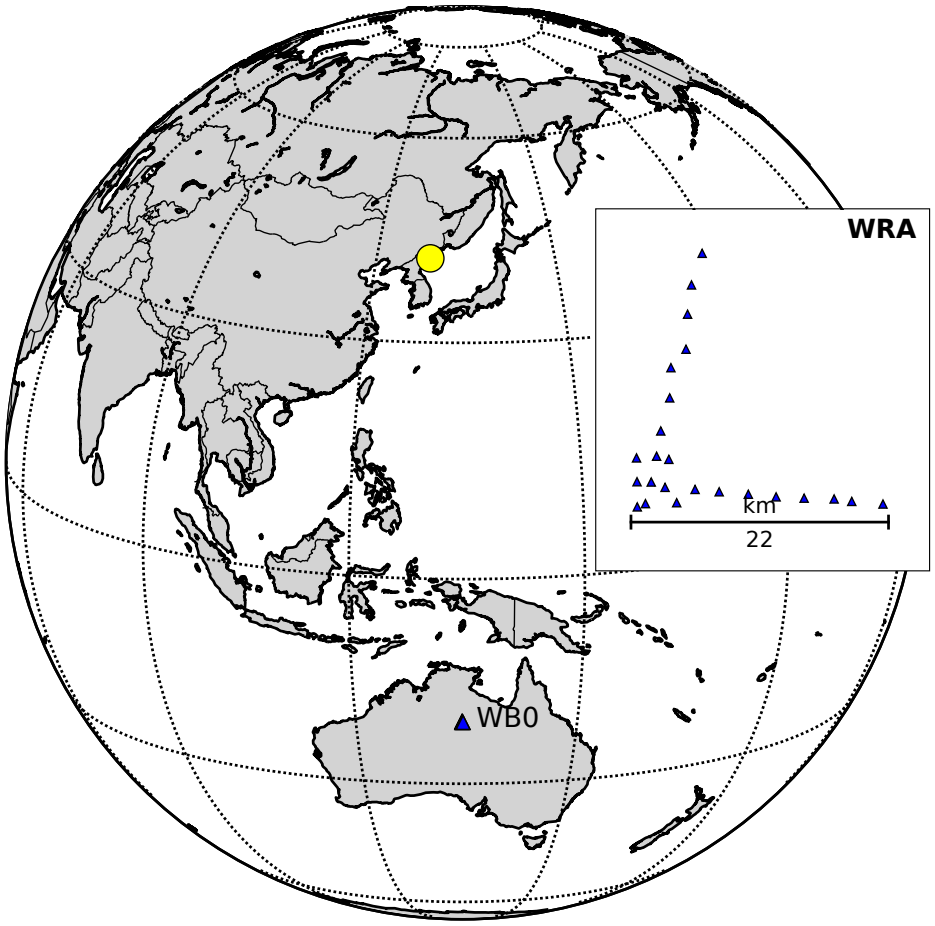

Figure 2.2: Location of the DPRK announced nuclear tests and the Warramunga Seismic Array (WRA) in Australia. Inset: array geometry of WRA.

Table 2.1: Backazimuth results for the DPRK 2006 nuclear test event

Array	$n = 1$	$n = 4$	$n = 8$
WRA	$353 \pm 5°$	$353 \pm 2°$	$353 \pm 1°$
MJAR	$302 \pm 5°$	$302 \pm 1°$	—
USArray	—	—	—

2.4 Adaptive versus non-adaptive

By seeking to directly align the waveforms in the traces, the adaptive stacking technique potentially offers much better signal improvement than techniques which are constrained by an *a priori* velocity model. The adaptive stack is able to adjust the relative time shifts as well as the relative amplitude of each station's contribution to the stack. However, this technique requires the waveforms to be similar enough between the stations that the misfit reaches a minimum

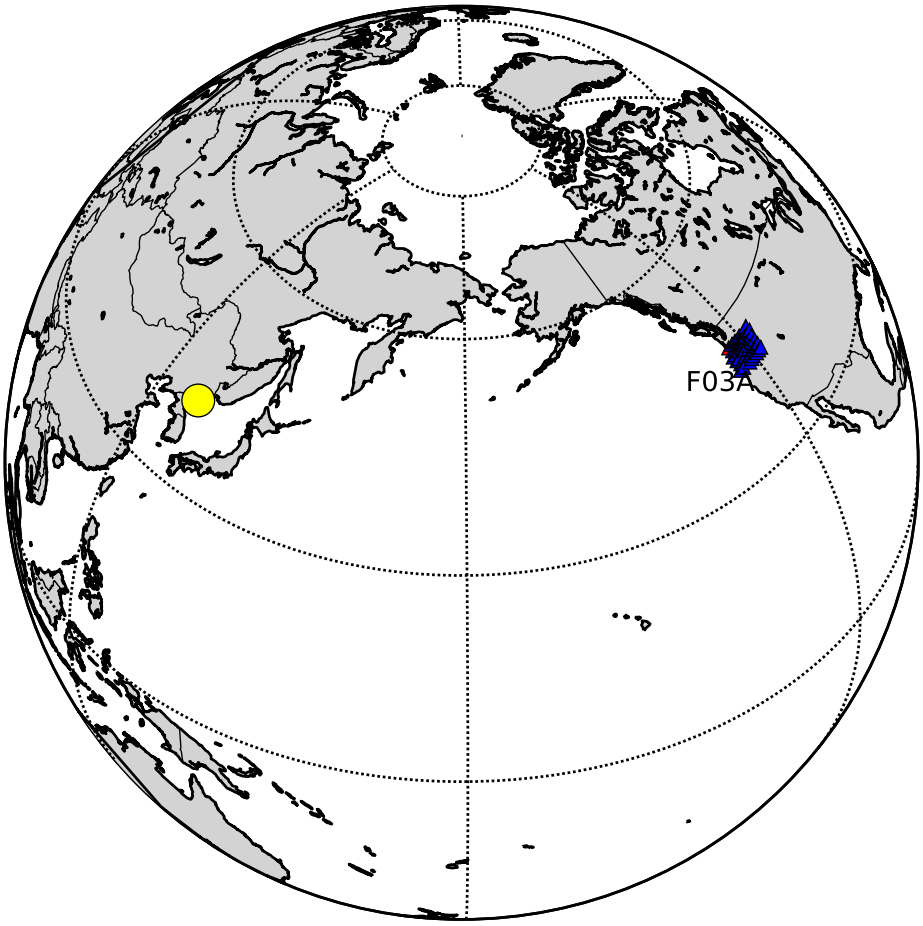

Figure 2.3: Location of the DPRK announced nuclear tests and a subset of the US Transportable Array (USArray) as of October 2006. The subset consists of 48 stations in the Pacific Northwest of the United States.

when the phase arrivals are aligned. In arrays where the shape of the arrivals are highly variable, the adaptive stack will not perform well.

I apply the adaptive stack method to data from nuclear tests conducted in North Korea from 2006 to 2016 at the Punggye-ri test site. Starting with the *ak135* shifts and running the `tcas` code on a 6 s window around the *ak135* predicted P arrival until the L_3 misfit is minimised stably over 10 iterations, I obtain new shifts for each trace in the stream.

The PSA array at Pilbara in Western Australia (Figure 2.7), is a 20 km aperture spiral arm array [47] deployed in a topographically homogeneous desert. In contrast the MJAR Matsushiro array in Honshu, Japan is deployed in an irregular arrangement over a mountainous region. MJAR is known to exhibit poor signal coherence [35] between the different stations of the array, while PSA is expected to show a high degree of coherence. In addition to this, MJAR is much closer to the test site (950 km), so the high-frequency seismic waves that are effectively filtered

out of the seismograms recorded at PSA are still present, and more sensitive to the local velocity structure of the array site. Figures 2.8 and 2.9 show seismic data for the 2013 announced nuclear test in North Korea recorded at each station of the arrays, alongside the array's correlation matrix and the semblance in the linear stack. At MJAR (Figure 2.8), the values across the correlation matrix are much lower than at PSA (Figure 2.9), indicating the dissimilarity of the waveforms between each element. Altering the frequency band to filter out higher frequencies showed no improvement to the correlation at MJAR, even down to 2 Hz, suggesting that this incoherence is not avoidable for body-wave arrivals from the DPRK test site. The semblance at PSA peaks during the expected arrival window of the P-phase, demonstrating good coherence between the signals. At MJAR, while the P-arrival can be seen in each trace, the semblance for the arrival window is low, due to the incoherent signal.

Both arrays record the first P-wave arrival from the event with good signal-to-noise ratio. However, while the waveforms recorded at PSA show high correlation and coherence across the array, the waveforms at MJAR are clearly dissimilar. This incoherency has a negative impact on the adaptive stacking method, which attempts to minimise the misfit between the traces. The shifts that result from the adaptive stacking procedure thus only delay the observed arrival of the P-wave, without improving the signal, as seen in Figure 2.11. At PSA the adaptive stack succeeds, but the resulting shifts are small and little improvement is made on the linear stack (see Figure 2.10), which already has a high SNR.

2.5 The effect of phase weighting

In Schimmel and Paulssen (1997) [69], the authors compare the phase-weighted stack (PWS) to the nth-root stack. They claim that the first-order phase weighted stack is expected to show "minimal" waveform distortion of coherent signals. However, as the order of the phase-weighting ν is increased, the waveform is distorted, with amplitudes from in-phase signals where the weighting is particularly high exaggerated with respect to other parts of the signal.

I apply the phase-weighted stacking procedure to data from the 2006 DPRK test recorded at WRA and MJAR, to test the SNR improvement for the smallest-magnitude event associated with nuclear testing in North Korea. Similar to the PSA array used in Section 2.4, WRA shows good coherence across the array elements, in contrast with the incoherent signals observed at MJAR. Locations of the arrays are pictured in Figures 2.1 and 2.2.

The data is filtered between 1–5 Hz to capture the body waves; for this event low frequency noise is present at WRA around 0.5 Hz. I calculate the PWS with $\nu = 1, 2, 4, 8$ using the *ak135* velocity model to predict the delay times.

At WRA (Figure 2.12), the first-order PWS slightly improves the SNR over the linear stack. However, higher-order PWS exaggerate the amplitude of a later phase arrival over the P-wave, creating a greatly distorted seismic waveform for the event.

At MJAR (Figure 2.13), increasing ν improves the amplitude of the P-wave arrival relative to the background noise, in contrast to the linear stack, where the P-wave is obscured by the noise. Despite the incoherence of the array, the phase-based method of weighting the signal succeeds where the best-fit approach of the adaptive stack does not. As at WRA, the higher-order PWS exaggerates later arrivals, but in this case without obscuring the first arrival.

Given the improved detection of the P-wave in the stacked traces at MJAR, I have produced phase-weighted vespagrams of order 1, 4 and 8 to compare with the nth-root vespagrams in Figure 2.4. The phase-weighted vespagrams are shown in Figure 2.14. The fourth- and eighth-order phase-weighted vespagrams show a smeared arrival at 10 s, coinciding with the expected P-wave arrival, but do not detect the P-wave itself. All the vespagrams show the exaggerated arrivals around 14 s, with the energy spread across a large range of backazimuths, suggesting it is an artefact of the phase-weighting process and not of physical origin. The PWS therefore does not show good results in the vespagram for this event.

2.6 Discussion

I will now discuss the results of the stacking techniques applied above at these arrays.

2.6.1 nth-root stacking

While the nth-root process distorts the shape of the waveform, the higher SNR increase compared with the linear stack is useful in signal detection, where the shape of the waveform itself is not important. Comparing power vespagrams for different orders of stack, I find that an n of 8 offers good signal detection properties, with good noise reduction and resolution of the slowness vector. The fourth-order stack has good noise reduction, but poorer resolution. Any higher-order stack outperforms the power vespagram for the linear ($n = 1$) stack. Given little difference in computation time for different n, the eighth-order stack is a good candidate for phase detection. Orders higher than 8 show little difference in noise reduction or slowness resolution.

2.6.2 Adaptive stacking

The results of the adaptive stacking procedure show that the technique performs well for arrays where the signal is coherent between the different stations of the array, and fails where it is incoherent, such as the MJAR array in Japan. For arrays where the local velocity structure differs little from a simple velocity model such as those used in constructing the linear stack, the signal improvement from using the adaptive stack is minimal, and the computationally faster linear stack will suffice. For arrays with observed velocity structure that differs significantly from a one-dimensional model, such as those for which I calculate slowness corrections in Chapter 4, the adaptive stack offers a different approach to improving the quality of the stacked waveforms.

It is possible to use the adaptive stacking technique, which changes the alignment of the traces in assembling the stack, with the phase-weighting procedure, which then weights the contribution from each station based on its phase. Provided the adaptive stacking is effective (i.e. the data is sufficiently coherent), the combination of both methods could yield a greater SNR improvement over either technique used on its own.

Coherence properties of arrays also depend on the particular frequency range of interest. For instance, waves with long periods of 10s of seconds would be expected to show good coherence even over a large array such as USArray, allowing the adaptive stack to be used.

2.6.3 Phase-weighted stacking

Phase-weighting the traces before stacking shows small or significant improvement to the SNR at regional and teleseismic distances. Increasing the order of the phase-weighted stack was found to improve the amplitude of the first arrival at the incoherent array MJAR. Compared with the linear and adaptive stack, the PWS achieves the best results at this array.

However, the use of the PWS for detection in e.g., a vespagram can be problematic, with distortion of the waveform leading to the exaggeration of some parts of the waveform at the expense of the first arrival. This effect was particularly extreme at WRA, where the nth-root stack is much more suitable for vespagram plotting.

2.6.4 Comparison of stacking techniques

Overall, the linear stack achieves good results at teleseismic arrays for events of magnitude 4.0–5.0 such as those from the announced nuclear tests in North Korea. At coherent arrays with good signal quality, little additional improvement is gained by performing the more computationally intensive nth-root, phase-weighted or adaptive processes.

For very large arrays, with instruments spread over hundreds of kilometres, these stacking techniques implemented with simple velocity models do not perform as effectively as at smaller arrays.

For event detection, the eighth-root stack achieves good signal-to-noise in a vespagram plot of the power in the stack.

At incoherent arrays, such as MJAR, the adaptive stack is unsuitable. Here the phase-weighted stack achieves the best results. Waveform distortion makes high-order PWS unsuitable for detection.

2.7 Conclusions

In this chapter I have described a number of time-series stacking techniques that can be applied to seismic data recorded at arrays. These include the linear stack, the nth-root stack, the phase-weighted stack, and the adaptive stack. I have also described an application of stacking in velocity

spectral analysis (VESPA), specifically the use of vespagrams to identify the backazimuth or slowness of incoming seismic phases. I applied these techniques to seismic data in order to identify the settings in which different stacking techniques perform well, and which order of stack to pick in the case of the nth-root process.

Based on these results, I recommend the following techniques for different studies, based on the arrays used. The focus is on the two principle aims of time-series stacking: improving the signal to study the properties of the seismic waveform, or to aid in the detection weak phases from small-magnitude events.

2.7.1 Signal improvement for waveform studies

For studies where the shape of the waveform must be preserved, the linear, adaptive and first-order phase-weighted stacks are suitable. Such studies might include correlations or convolutions of waveforms in receiver function studies, or in calculating seismic amplitude or magnitude ratios.

Small aperture arrays at teleseismic distances generally show good coherence properties, whereas for very large arrays, or those much closer to the source, the coherence of the wavefield recorded at the array is worse.

- For arrays with good coherence, the adaptive stack yields the best results. It can also be combined with the first-order phase-weighted stack without distorting the shape of the waveform.
- If the data is incoherent between the stations of the array, the first-order phase-weighted stack should be used alone.
- For very large arrays, the adaptive stacking procedure will perform well for low-frequency phases, but for higher frequency waves the plane-wave approximation fails, and none of the above stacking techniques is suitable.

2.7.2 Detection of weak phases

If the shape of the waveform is not important, and maximal signal-to-noise improvement is required, the nth-root stack performs well.

- For arrays with good coherence, the eighth-root stack performs well. Higher orders of nth-root stack have little effect, and phase-weighting yields little additional improvement and risks amplifying spurious signals.
- If the data is incoherent between the stations of the array, the eighth-order phase-weighted stack shows an improvement over the nth-root stack.

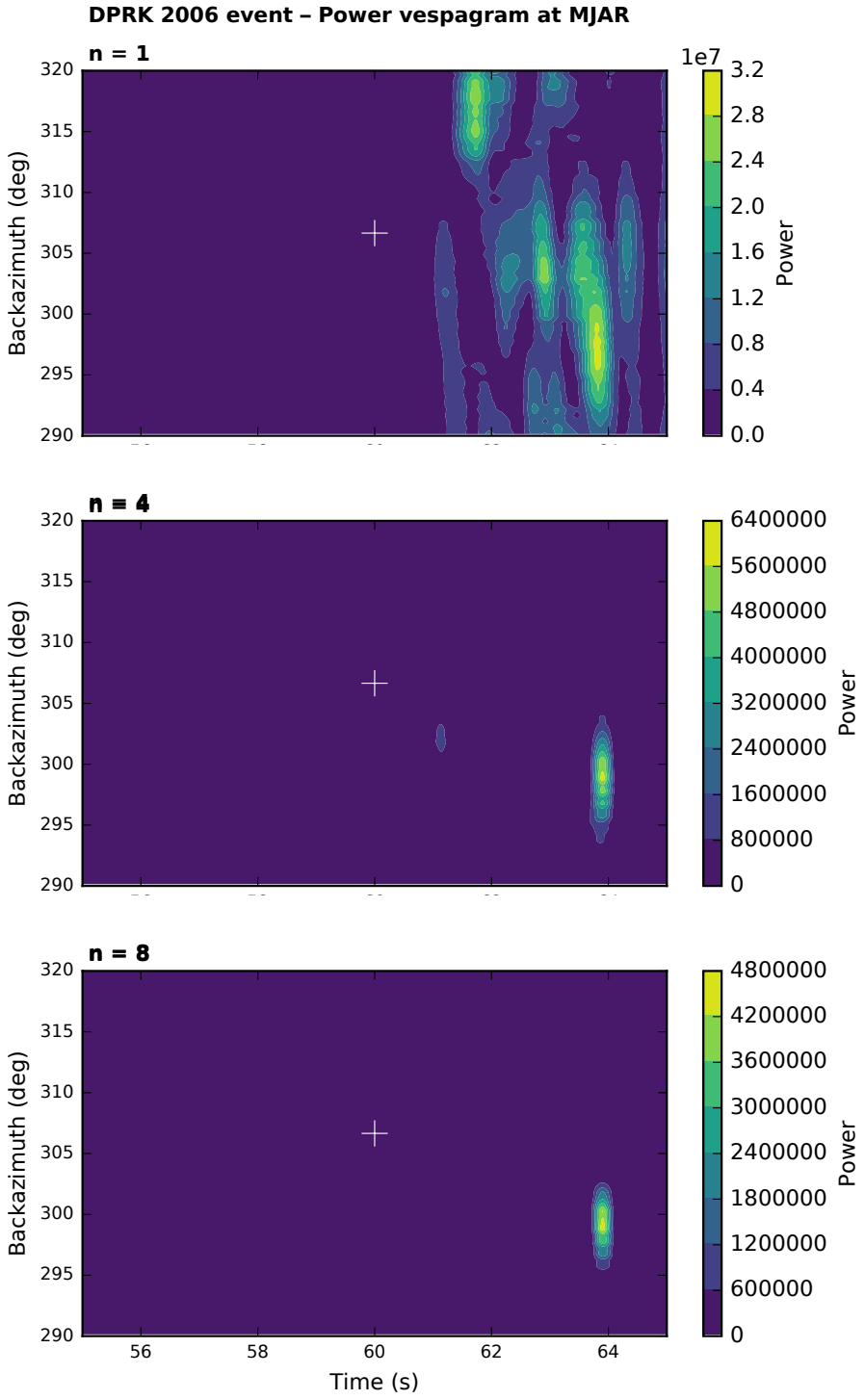

Figure 2.4: Backazimuth vespagrams for the DPRK 2006 announced nuclear test recorded at MJAR. The power in the nth-root stack is shown, with $n = 1, 4, 8$ from top to bottom. The predicted P phase arrival at the array reference station MJA0 is marked with a white cross. The linear stack has too much noise to clearly resolve the P-wave. The 4th-root stacks reduce this noise, showing an arrival around 1 s after the predicted arrival with a smaller backazimuth. However in the 8th-root stack the P-wave is obscured by a higher-amplitude arrival at around 14 s.

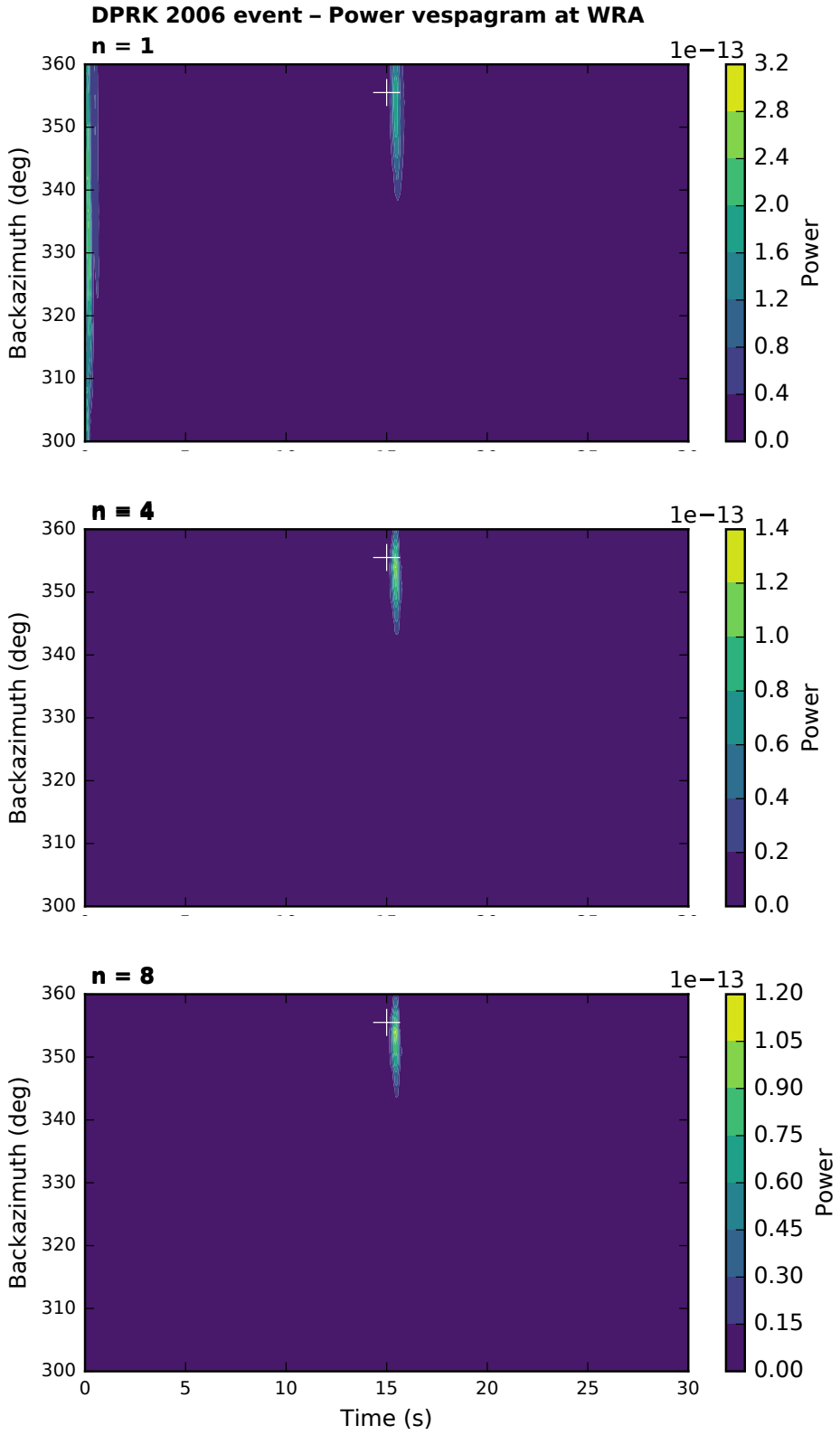

Figure 2.5: Backazimuth vespagrams for the DPRK 2006 announced nuclear test recorded at WRA. The power in the nth-root stack is shown, with $n = 1, 4, 8$ from top to bottom. The predicted P-wave arrival is at 17 s with a backazimuth of 355° (white cross) from the array reference station WB0. The P-wave is clear for all n, with a slightly higher backazimuthal resolution with higher n.

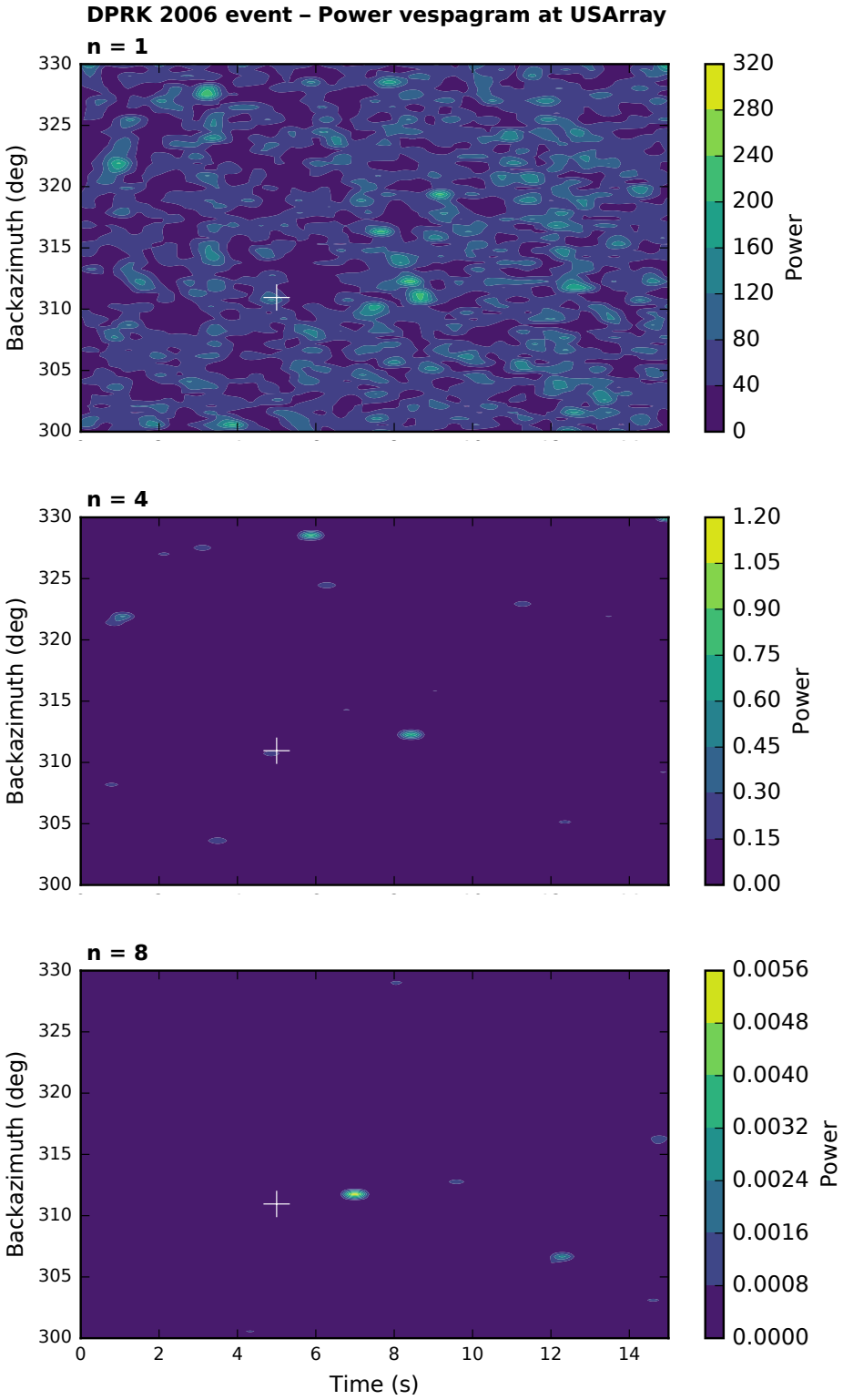

Figure 2.6: Backazimuth vespagrams for the DPRK 2006 announced nuclear test recorded at USArray. The power in the nth-root stack is shown, with $n = 1, 4, 8$ from top to bottom. The theoretical P-wave arrival at the array reference point is marked with a white cross. It is not clear whether the P-wave is resolved, or whether the power maxima are due to errors due to the non planarity of the wavefront across the very large array. The 8th-root vespagram shows a clear phase arrival at the correct backazimuth around 2 s later than predicted. This later apparent arrival could be due to non-planarity of the arrival across the large array.

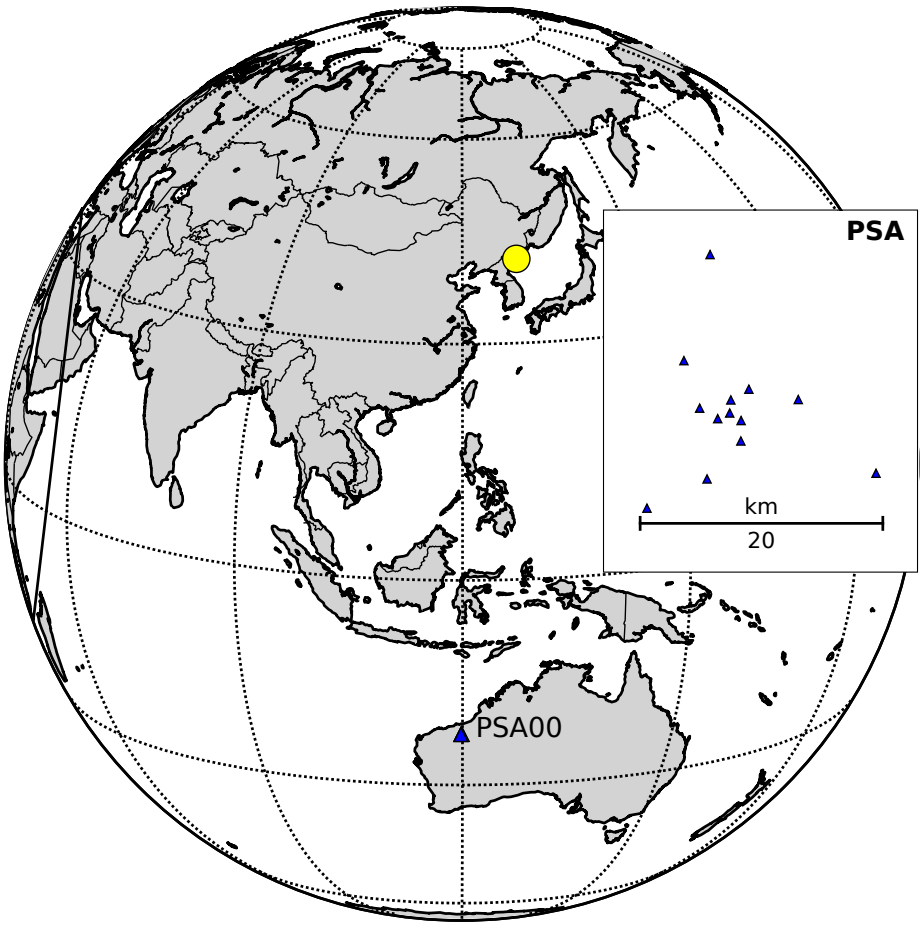

Figure 2.7: Location of the DPRK announced nuclear tests and the Pilbara Seismic Array (PSA) in Australia. Inset: array geometry of PSA.

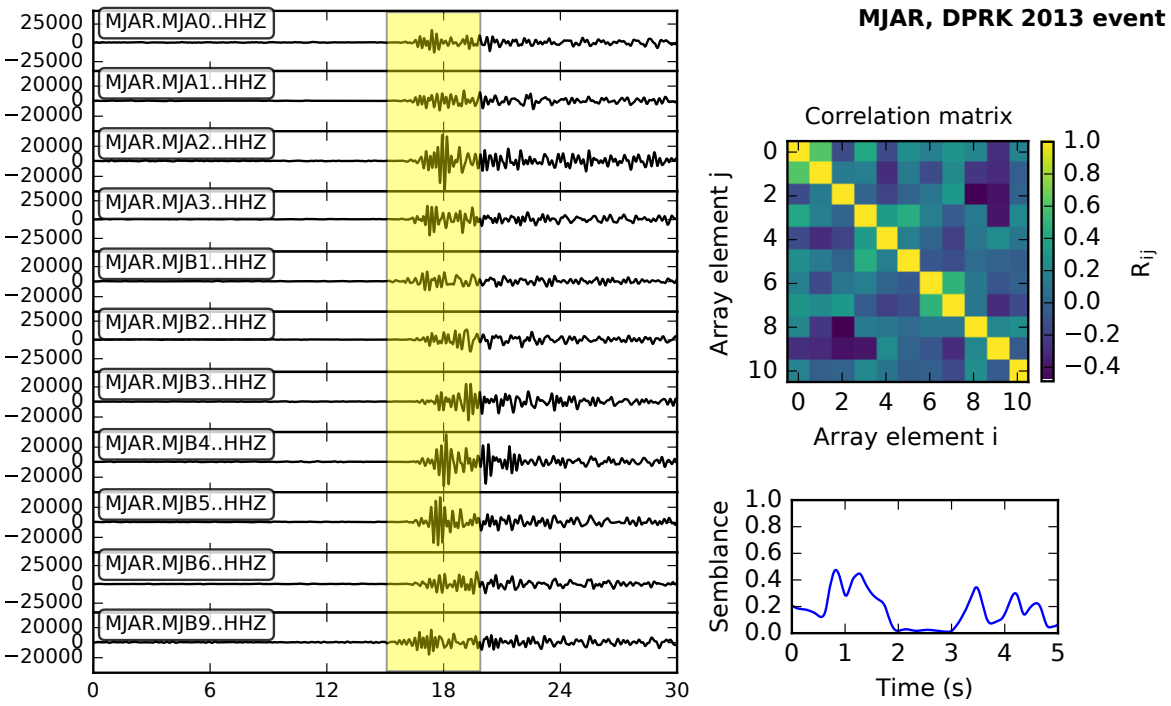

Figure 2.8: Adaptive stacking at the MJAR array. The individual traces for each station in the array are plotted to the left, filtered between 1 and 5 Hz. The yellow box shows the time window used in the adaptive stacking program. The correlation matrix and the semblance show a much lesser degree of correlation and coherency between the array elements than is observed at PSA. This is also visible in the dissimilarity between the waveforms themselves.

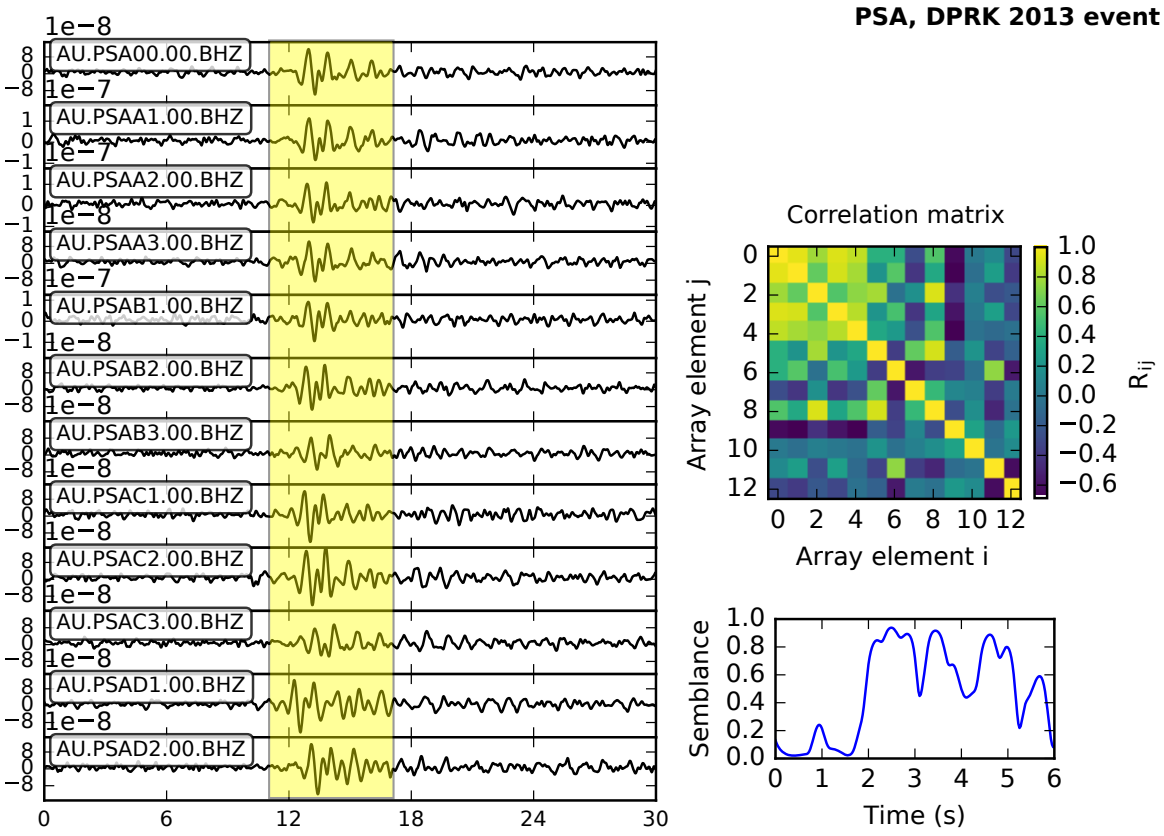

Figure 2.9: Adaptive stacking at the PSA array. The individual traces for each station in the array are plotted to the left, filtered between 1 and 5 Hz. The yellow box shows the time window used in the adaptive stacking program. To the right is the correlation matrix for the time window, showing good correlation between many array elements. Beneath is a semblance plot for the same window, with the semblance calculated in a 0.5 s time window in the linear stack. The semblance for the P-arrival is high, demonstrating the high coherence of the signal recorded at PSA.

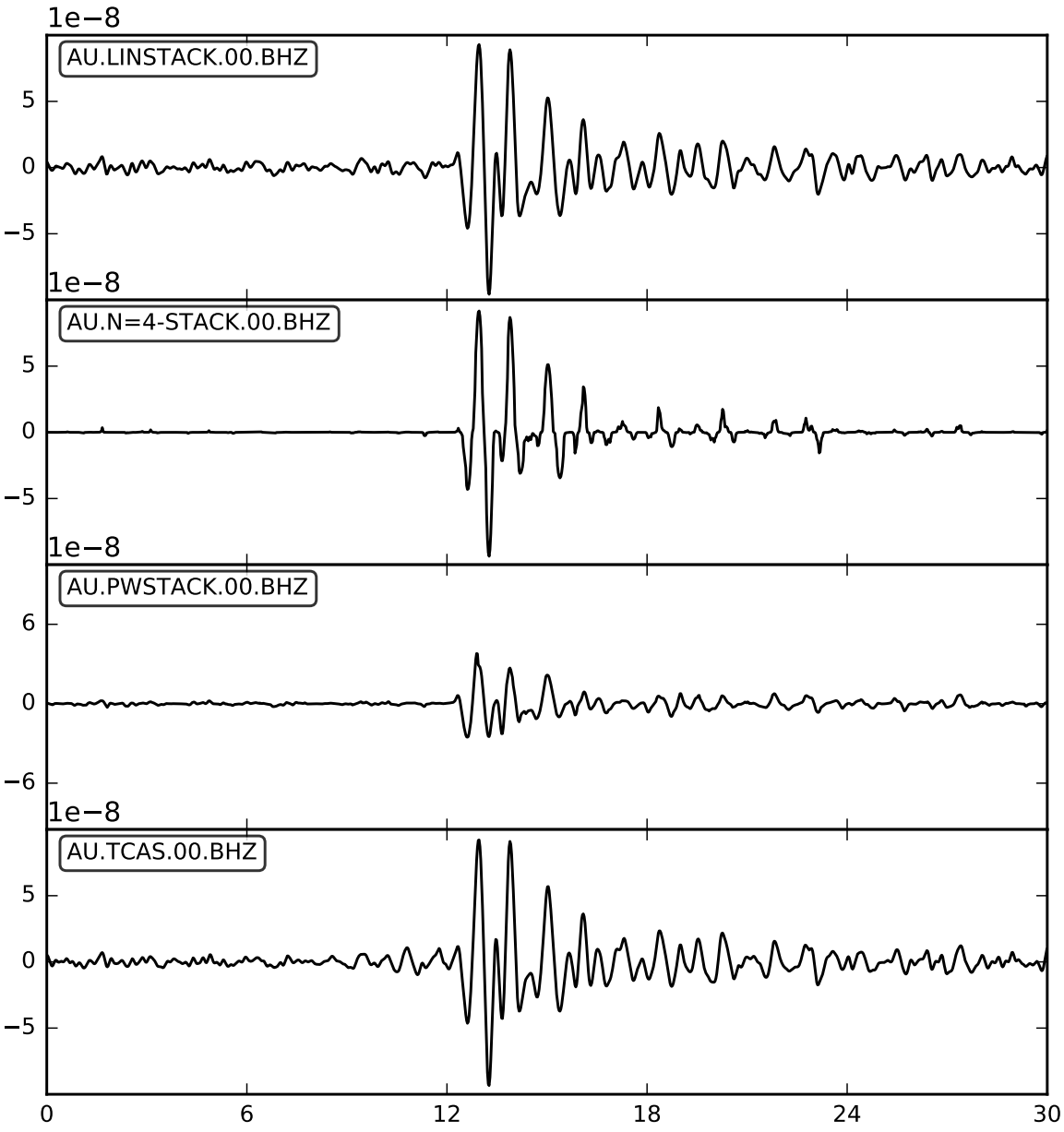

Figure 2.10: Stacked waveforms for the PSA array. From top to bottom, linear stack, 4th-order nth-root stack, first-order phase-weighted stack and adaptive stack. The predicted arrival is at 12 s. The adaptive time shifts make little change to those predicted by the velocity model, but the linear and adaptive stacks both have the best SNR without distorting the waveform, as in the case of the 4th root stack.

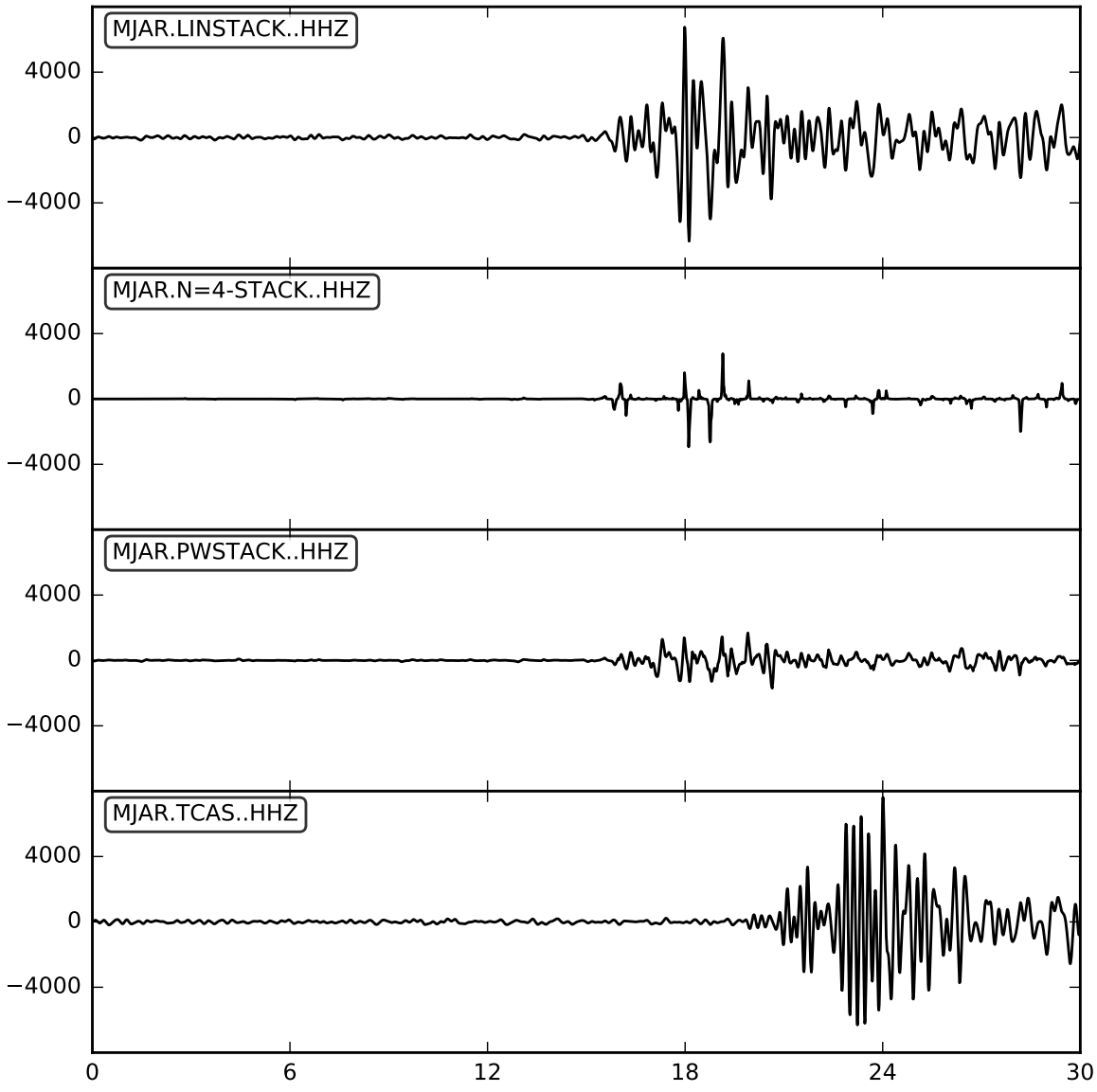

Figure 2.11: Stacked waveforms for the MJAR array. From top to bottom, linear stack, 4th-order nth-root stack, first-order phase-weighted stack and adaptive stack. The adaptive stack has failed to resolve the P-arrival due to the low waveform similarity between the array elements.

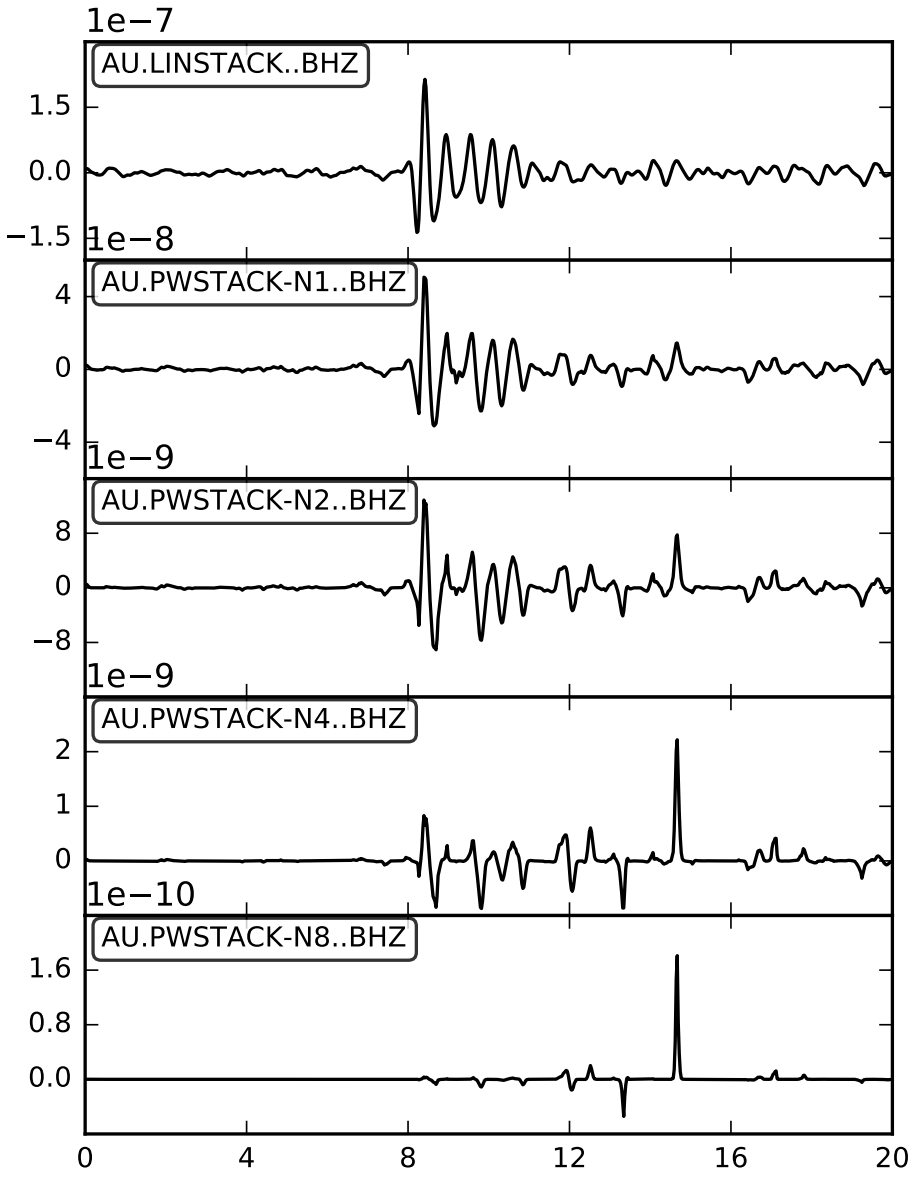

Figure 2.12: Phase-weighted stacks for WRA, with the linear stack (top) for comparison. The order of the phase-weighting is increased through $\nu = 1, 2, 4, 8$. The first-order (and higher) phase-weighted stacks show slight noise reduction compared with the linear stack. However, as ν is increased beyond 1, the stacking exaggerates a later seismic arrival around 14 s, and the amplitude of the P-wave (at 10 s, the theoretical arrival time) is diminished. This makes the higher-order PWS particularly unsuitable for phase detection.

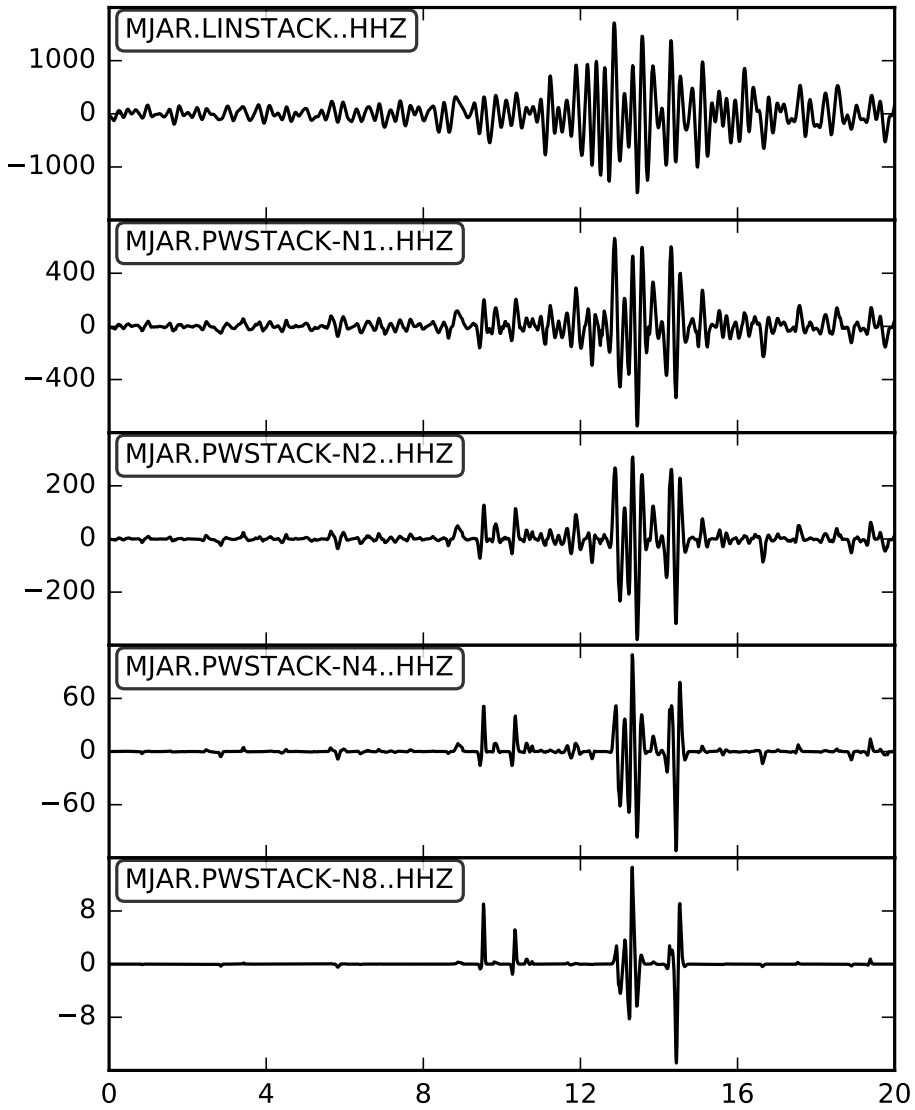

Figure 2.13: Phase-weighted stacks for MJAR, with the linear stack (top) for comparison. The PWS is effective for all ν in distinguishing the P-wave around the theoretical arrival at 10 s, which has an amplitude little above the noise level in the linear stack. As at WRA, increasing ν exaggerates later arrivals to a greater degree, but at MJAR the P-wave is not lost even in the eighth-order PWS.

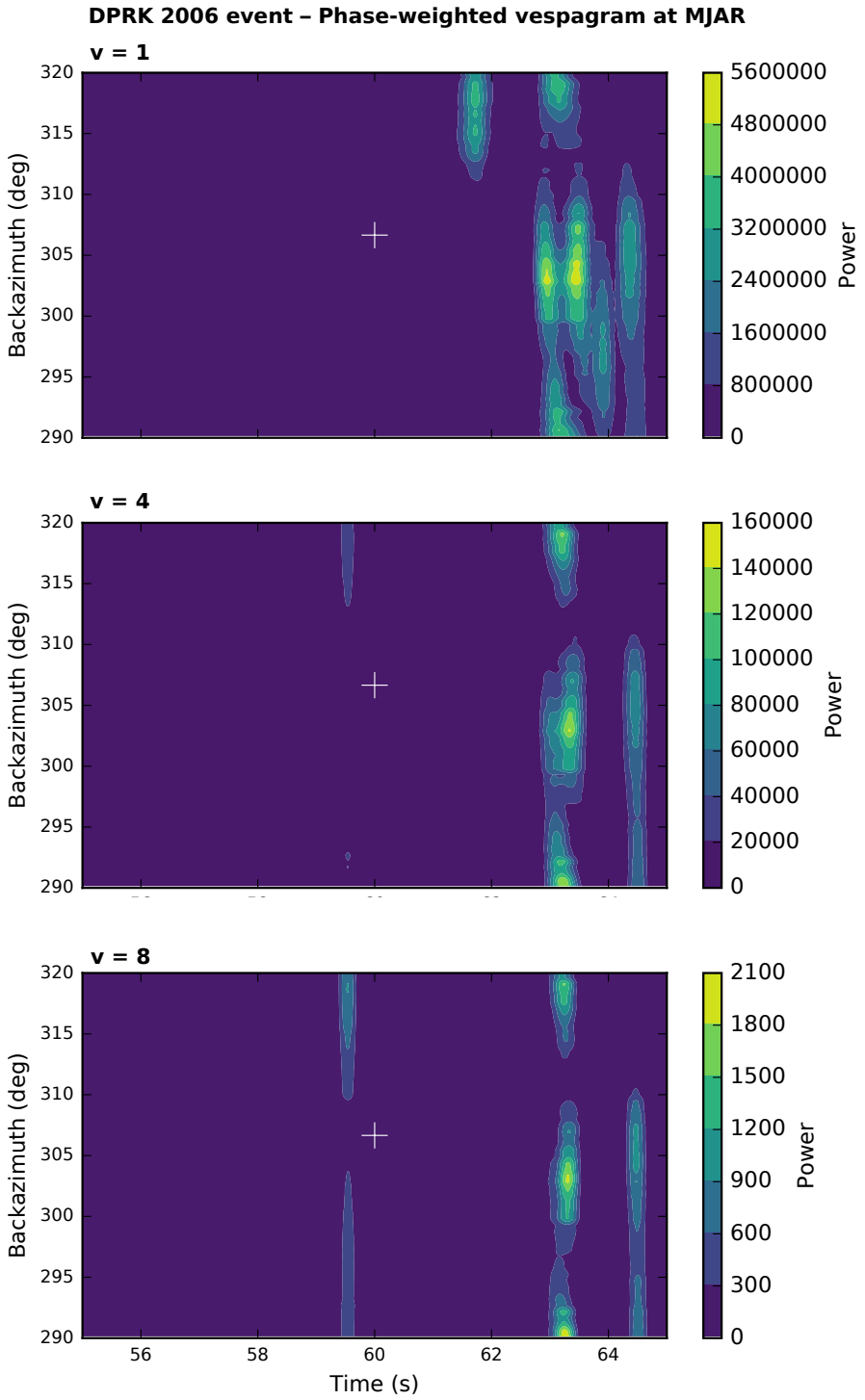

Figure 2.14: Phase-weighted vespagrams for the 2006 DPRK event recorded at MJAR. The predicted P arrival is marked with a white cross. All orders of phase-weighted vespagram fail to resolve the P-arrival. There are local power maxima at the predicted time, but it is smeared over a range of backazimuths away from the source at 307°. The phase-weighting offers no benefit to the nth-root vespagrams in Figure 2.4.

CHAPTER 3

VesPy: a Python package for array analysis

An increasing number of seismologists, among other scientists, are using the Python programming language as their day-to-day software for data collection, analysis and visualisation. Python offers a number of advantages compared to other programming languages. It is cross-platform and open-source, available for individuals and institutions for free. It is an interpreted language, capable of running line-by-line without separately compiling beforehand. This allows for an experimental approach to running scripts and software development useful in data analysis. It emphasises a human-readable approach to coding that makes it a good language for students and early-career researchers looking to program for the first time. It is also extensively customisable, with a large online library of packages that add functions and data structures useful for scientific computing. For example, the Scientific Python (SciPy) project[1] develops and maintains popular scientific computing modules in Python. NumPy contains numerical functions and a mutli-dimensional array-based data structure. Matplotlib provides graphing and data plotting functions similar to those in the commercial MATLAB software, but for free. Pandas offers a flexible database structure for storing data of different types.

In seismology, the ObsPy package contains several useful modules for reading seismographic data, data structures for handling catalogues of events and inventories of stations, geodetic functions for calculating great-circle geometries, FDSN functions for downloading seismic data from the internet, and an implementation of the TauPy ray-pathing software. ObsPy is thus very useful for handling and analysing seismic data.

However, ObsPy lacks native support or implementation of various array analysis methods. My motivation for this part of the project was to develop an open-source software package, VesPy, in Python, that would enable seismologists to perform common array analysis techniques, while

[1] https://docs.scipy.org/doc/

taking advantage of ObsPy's handling of different seismic data formats.

This chapter contains a description of VesPy and three case studies used to illustrate the package's utility.

3.1 Introduction

The development of modern seismology has seen a considerable increase in the availability of data and new analysis techniques arising from the increased global deployment of networked arrays of digital seismometers that allow for the application of array signal processing methods ([66]). One key method that can be applied to seismic array data is stacking, also known as beamforming. Time series data from each instrument within the array are combined, taking into account the time delays due to the moveout of the wavefield between instruments. The desired signal propagates coherently across the array, and stacks constructively so that its amplitude is enhanced with respect to the noise, which is incoherent between the different receivers. Thus, the stacked data can reveal weaker signals from lower-amplitude phases and lower-magnitude events than can be observed in the individual seismic traces.

The first large aperture seismic arrays were deployed in the 1960s at a time when high-quality data was sought, among other reasons, for the purpose of monitoring nuclear weapons testing, known as forensic seismology [12]. The early seismic arrays often had a cross-shaped UKAEA geometry, such as the YKA array in Yellowknife, Canada, or ESK in Eskdalemuir, Scotland. For better azimuthal resolution, modern arrays are frequently designed with concentric circular geometries, such as ARCES in Norway [71]. Kennett *et al.* (2015) [47] suggest that similar results are possible with many fewer instruments using spiral-arm geometries. Arrays for global seismology and nuclear monitoring have apertures on the order of 10 km, with smaller, denser arrays having better angular resolution at the cost of slowness resolution. Seismic arrays are also frequently used in exploration geophysics, involving much larger numbers of sensors over a much smaller area, designed to pick up very low magnitude microseismic events close to the array (e.g. [51]). Array processing methodologies are not necessarily dependent on the scale of the geometry of the arrays involved, and can be effective both with small and large numbers of instruments and apertures, so long as the propagation of the wavefronts across the array is modelled accurately.

In this chapter, I present a software package, VesPy, written in Python, that provides functions for performing various seismic array analysis methods. I was motivated to create a package that will work with existing seismological toolkits in Python, particularly the popular ObsPy package [4], and to produce outputs in NumPy [81] and Matplotlib [40] so that the analysis done with this package can easily be incorporated into wider seismological studies carried out in Python.

3.2 Methodology

3.2.1 Array methods in seismology

Seismic arrays consist of multiple seismic sensors (whether single-component geophones or three-component broadband seismometers) all reporting their data to the same network. By providing digital seismic time series data with common time stamps, array data is ideal for combining data from different instruments for the same source event.

Beamforming refers to the process of creating a beam, i.e. a signal that represents the incoming wavefield (e.g. from a low magnitude seismic tremor) with a minimum of background noise. The same principle useful in, for example, radar antenna design is used in seismology. The beam is constructed as a delay-and-sum stack, where the seismic waveforms at each instrument are shifted by a delay time to account for the propagation of the wavefield across the array, and then summed together. If the targeted phase is a plane wave, and its arrival is aligned across all stations, the stack will amplify the signal while suppressing noise (generally signals from unwanted microseismic sources near the receiver).

There are various stacking methodologies implemented in this package, ranging from the simple linear stack, to more intensive statistical methods.

3.2.2 Time-domain stacking

When stacking the waveforms from each receiver in the array, it is necessary to first shift each trace such that the desired seismic phase is aligned, and the signals sum constructively. The delay times τ_j are given by

$$\tau_j = \vec{r}_j \cdot \vec{s} \tag{3.1}$$

for receivers at location $\vec{r}_j$ with respect to a common reference point (often the central element of the array) and an incident wavefield propagating across the array with horizontal slowness vector $\vec{s}$. The slowness vector depends on the velocity model assumed.

The delay times are then applied to the seismic trace $u_j(t)$ (the amplitude recorded at each station at time t) by shifting the time series to $u_j(t-\tau_j)$.

3.2.2.1 Linear stack

The linear stacking procedure (also known as delay-and-sum) simply finds the mean amplitude of the shifted traces across the whole array as a function of time. The linear stack as a function of time is given by

$$v(t) = \frac{1}{N}\sum_{j=1}^{N} u_j(t-\tau_j), \tag{3.2}$$

where N is the total number of instruments within the array. For the ideal case where there is only one coherent signal in a background of random, incoherent noise, the signal-to-noise ratio (SNR) of the linear stack is improved by $\sqrt{N}$ with respect to the unstacked traces.

3.2.2.2 nth-root stack

The nth-root stack ([43]) is calculated in two steps. In the first, the nth root of each of the seismic traces is taken before they are summed:

$$v'_n(t) = \frac{1}{N}\sum_{j=1}^{N} u_j(t-\tau_j)^{\frac{1}{n}}\,\mathrm{sign}\,[u_j(t-\tau_j)], \tag{3.3}$$

where $n > 1$ is an integer. The result is then raised to the nth power to produce the stack

$$v_n(t) = \left|v'_n(t)\right|^n \mathrm{sign}\,[u_j(t-\tau_j)]. \tag{3.4}$$

Taking the root of the data reduces the difference in amplitude across the traces, diminishing the relative importance of high amplitude segments of the data and making the coherence of the signal across the array contribute more to the resulting stack. The nth-root stack therefore achieves better suppression of incoherent noise compared with the linear stack, at the cost of preserving the shape of the waveform. The nth-root stack is therefore not suitable for waveform studies, beyond polarisation.

A larger n increases the sharpness of the phase arrivals, but results in more distortion of the waveform. Common choices of n are 2 and 8, with [43] finding that the 8th-root stack shows a good approximation of the signal, with little further improvement for $n > 8$. Four nth-root stacks for the same event with different n are shown in Figure 3.1.

3.2.2.3 Phase-weighted stack

The phase-weighted stack ([69]) uses a direct measurement of the instantaneous phase of the seismic traces at each station to weight the contribution of that station to the stack. The phase at each station $\phi_j(t)$ is calculated from the complex traces as follows:

$$\phi_j(t) = \arg[u_j(t) + iH(u_j)(t)], \tag{3.5}$$

where $H(u_j)(t)$ is the Hilbert transform of the seismic trace $u_j(t)$.

The phase is then used to weight the contribution of each station to the stack

$$v_\phi(t) = \frac{1}{N}\sum_{j=1}^{N} u_j(t-\tau_j)\left|\frac{1}{N}\sum_{j=1}^{N} e^{i\phi_j(t-\tau_j)}\right|^\nu, \tag{3.6}$$

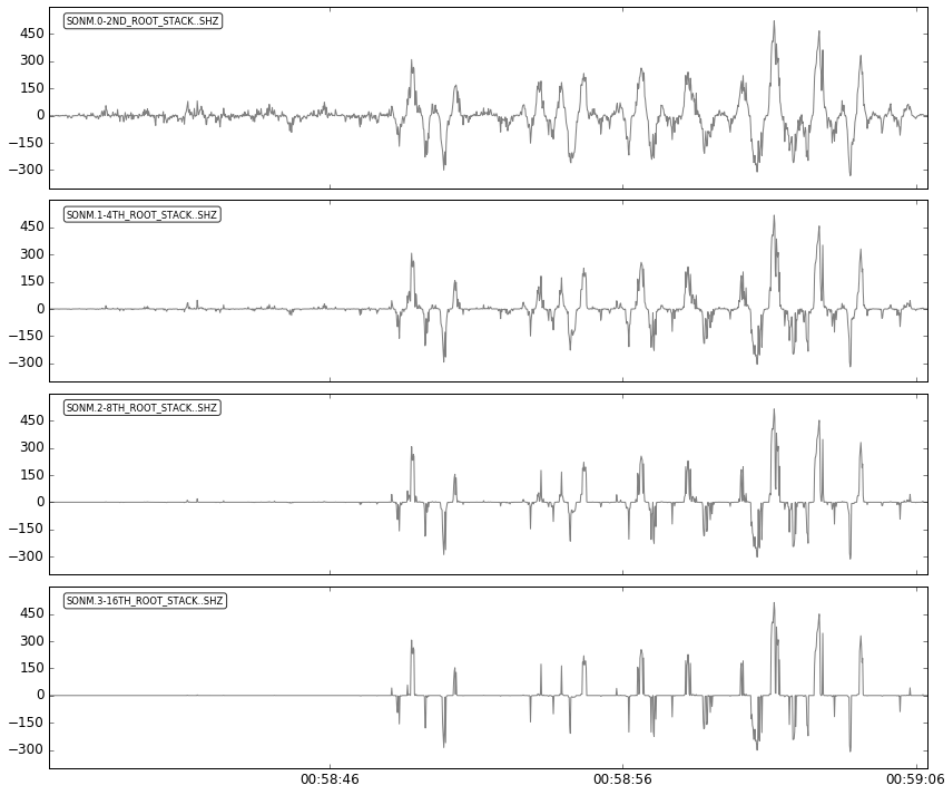

Figure 3.1: Comparison of nth-root stacks with differing n for the 2009 DPRK nuclear test recorded at SONM. From top to bottom, $n = 2, 4, 8, 16$. As n is increased, the signal becomes sharper. This improves the stack as a detector of individual phases, but more of the waveform shape is lost.

where $\nu \geq 1$ is an integer controlling the sharpness of the phase-weighting that is applied. The phase-weighted stack results in a better noise reduction than the linear stack, but due to the nonlinear weighting, the shape of the waveform is not preserved (similar to the nth-root stack). A comparison of how the phase-weighted stack changes with different ν is shown in Figure 3.2. As with the nth-root stack, a higher order weighting leads to sharper phase arrivals, with more loss of the signal waveform. The first-order phase weighting gives a better approximation of the signal waveform, but the 8th-order phase-weighted stack is more useful if only the arrival time of the targeted phases is desired.

A comparison of the three stacking techniques above, as implemented in VesPy, is shown in Figure 3.3. In arrays with a large number of elements, where the linear stack provides a good SNR improvement, it is generally the best choice, both because it is computationally quicker, but also because it preserves the shape of the waveform, allowing the stack to be useful for a greater number of applications. Where the linear stack fails, the phase-weighted stack approximates the signal better than the nth-root stacks, but is the most computationally intensive of the stacking methods. In situations where the full waveform is not needed, the 8th-root stack gives sharp

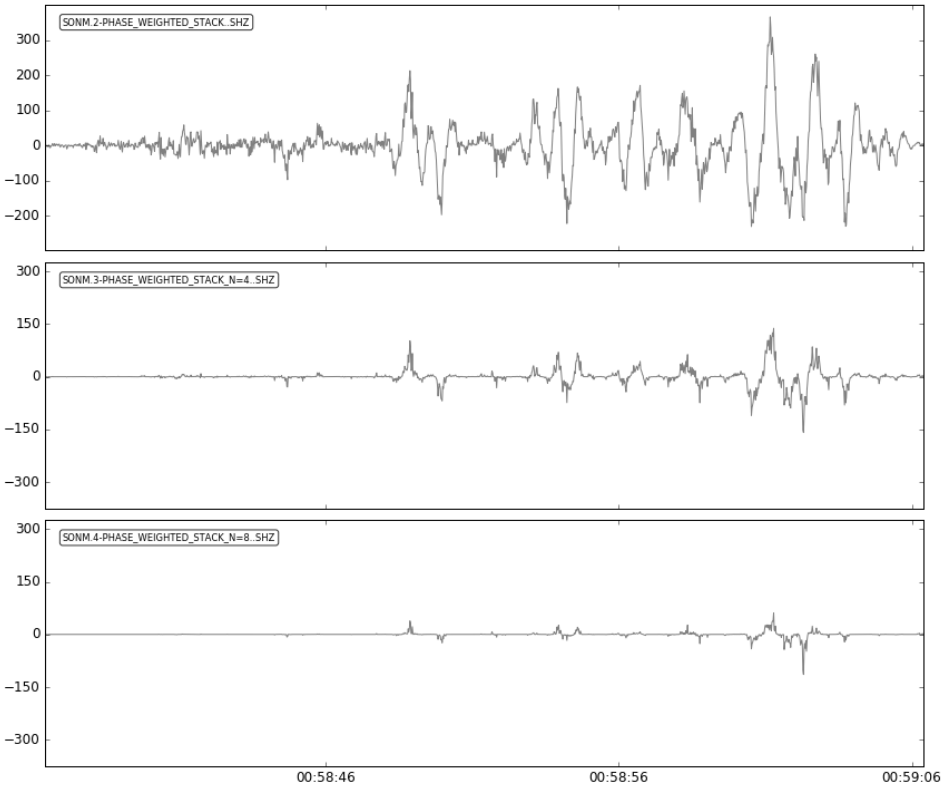

Figure 3.2: Comparison of phase-weighted stacks with differing weightings ν for the 2009 DPRK nuclear test recorded at SONM. From top to bottom, $\nu = 1, 4, 8$. The 1st-order phase-weighted stack gives good SNR improvement over the linear stack (see 3.3). Higher order stacks decrease the amplitude of the noise further, but also diminish the amplitude of the signal, distorting the shape of the waveform.

phase arrivals that are useful for phase detection in vespagrams, as shown in Section 3.4.1.

3.2.3 Semblance and F-statistic

In addition to the stacked waveforms, the detection of seismic phases can also be assisted by using statistical methods such as the semblance and the F-statistic.

The semblance, defined by Neidell and Taner (1971) [54], is a measurement of the power in the stack, normalised by the total power in the unstacked traces.

$$S(t) = \frac{\left(\sum_{j=1}^{N} u_j(t)\right)^2}{N \sum_{j=1}^{N} u_j^2(t)}. \tag{3.7}$$

The semblance $S(t)$ thus provides a measurement of the coherency of the signal across the array, and is commonly used for velocity analysis in seismic reflection surveys (e.g., [32]), and detecting microseismic events (e.g., [75]).

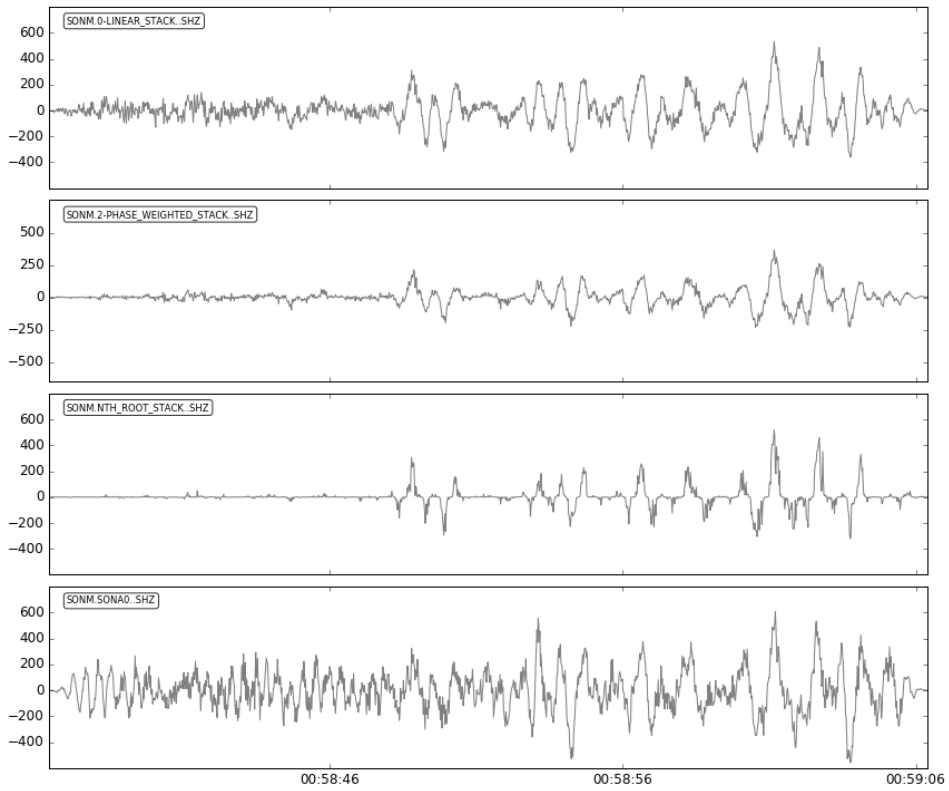

Figure 3.3: Example output, plotted as ObsPy traces, for the various VesPy stacking methods for seismic data from the announced underground nuclear test in North Korea on 25th May 2009 recorded at the IMS primary seismic array SONM. The nth-root stack has $n = 4$. The unstacked waveform from one of the array stations is show at the bottom for reference. The SNR is improved for the linear and phase-weighted stacks. The 4th-root stack shows clear spiky arrivals above the background noise, but the shape of the waveform is not preserved.

The F-statistic, first formulated by Blandford (1974) [7] provides an estimate of the instantaneous SNR. Here I calculate it over a short time window as

$$F(t) = \frac{N(N-1)v(t)^2}{\sum_{j=1}^{N} u_j(t-\tau_j)^2 - Nv(t)^2}, \tag{3.8}$$

where $v(t)$ is the amplitude of the stacked waveform at time t (here the linear stack, as in Equation 3.2). The squared amplitude of the stack is used as an estimate of the signal power, and is subtracted from the amplitude in each trace $u_j(t)$ as an estimate of the noise.

This formulation of F assumes the noise is uncorrelated between the stations in the array, although a more generalised F-detector has been formulated by Selby (2008) [72] to work at arrays with correlated noise.

A comparison of the semblance and F-statistic as implemented in VesPy is shown in Figure 3.4 (a more detailed comparison can be found in Chapter 2). Both $F(t)$ and $S(t)$ can be used as signal

detectors, with a detection being made when either statistic is higher than a determined threshold. $F(t)$ in particular can be more conveniently converted into a probability of detection. It is therefore suited to event detection and assessing the probability of a false alarm in forensic seismology. Meanwhile, the semblance can be used as a weighting factor in assembling semblance stacks, which are frequently used for event detection in studies of microseismicity.

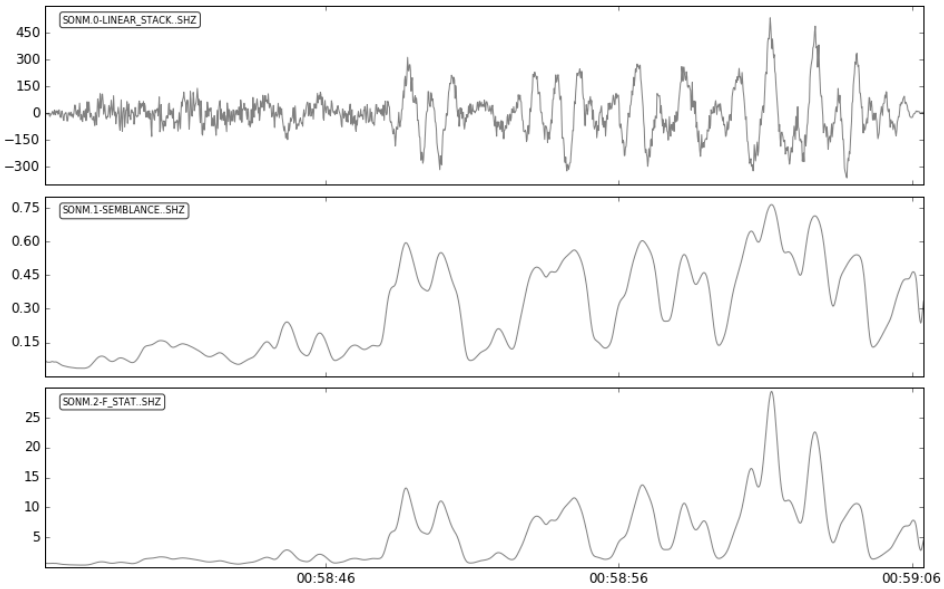

Figure 3.4: Example output, plotted as ObsPy traces, of the semblance and F-statistic methods for the 2009 DPRK nuclear test recorded at SONM. The linear stack for the same event is shown above for reference. The semblance measures the coherence of the signal in the stack, and the F-statistic the SNR. The F-statistic in particular gives a good indication of one the phases arrive, with a sudden increase in F above the background noise.

3.2.4 Velocity spectral analysis

When the slowness or backazimuth of an incident wavefield is not known, stacks for a range of slownesses can be calculated and the amplitude or (more frequently) the power in the stack plotted together as a function of time and slowness in a plot called a vespagram [22]. An equivalent vespagram can also be calculated for a fixed scalar slowness and varying backazimuth. The power in the nth-root stack or the F-statistic can be used to give a sharper phase arrival in the vespagram, allowing the apparent slowness or backazimuth of the seismic phase to be estimated with greater precision.

Example VesPy vespagram plots calculated using the power in the linear stack, the power in the 4th-root stack, and the F-statistic in the linear stack are shown in Figure 3.5. Vespagrams are useful for detecting and comparing the slowness and arrival times of multiple phase arrivals from a particular seismic event, or multiple events occurring within a short time of each other.

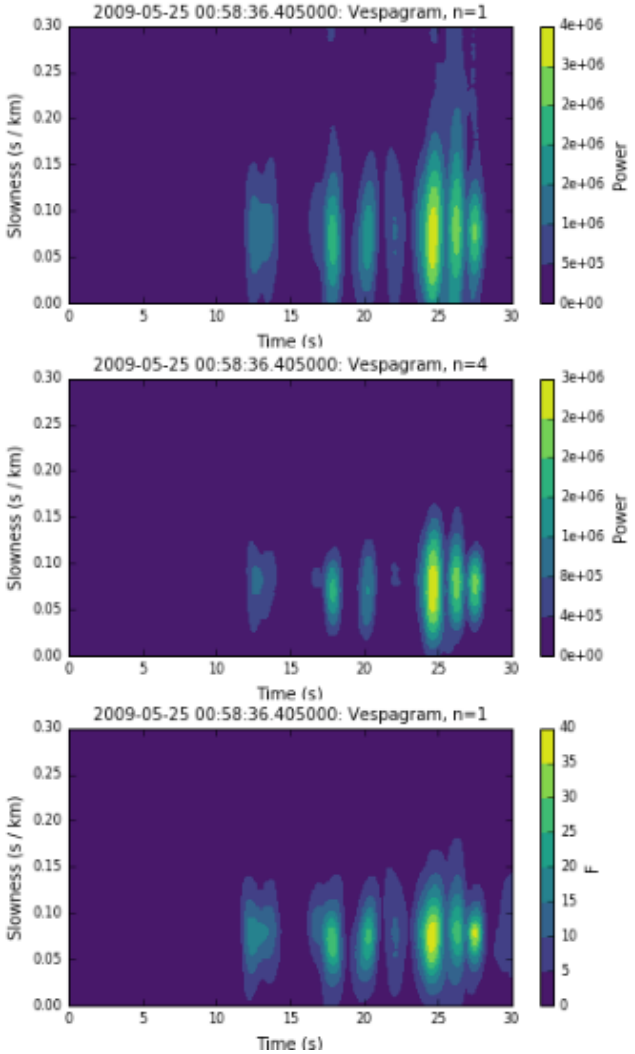

Figure 3.5: VesPy vespagram plots for the 2009 DPRK nuclear test recorded at SONM. From top to bottom: power vespagram for the linear stack, power vespagram for the nth root stack (with $n = 4$), and F-statistic vespagram for the linear stack. The higher-order power vespagram and the F-statistic give a better slowness resolution when identifying the phase arrivals. For this event, the first P-wave arrival can be identified with a slowness of 0.8 s/km in the 4th-order vespagram, but the same arrival is spread across the range 0.04–0.11 in the first-order plot.

3.2.5 Frequency-wavenumber analysis

Frequency-wavenumber (fk) analysis, developed by Capon (1969) [14], is performed in the frequency domain and provides a computationally quick way to estimate the complete two-dimensional slowness vector (i.e. the scalar slowness and backazimuth) of the signal. For an array comprising N instruments, a square slowness grid of size $N \times N$ is generated for the slowness range of interest in the north-south (s_y) and east-west (s_x) directions. The fast Fourier transform (FFT) of the seismic traces recorded at each station is calculated. The fk beam at each point in

the slowness grid is then calculated, by analogy with the time-domain beam (Equation 3.2), as

$$V_{mn}(f) = \frac{1}{N}\sum_{j=1}^{N} e^{-2\pi i \tau_{mn} \otimes f} U_j(f), \tag{3.9}$$

where $U_j(f)$ is the amplitude of the Fourier spectrum at the discrete sample frequencies f for the jth station in the array, and $\tau_{j,mn} = s_{x,n} r_{x,j} + s_{y,m} r_{y,j}$ is the corresponding delay time at the station for the point on the slowness grid $(s_{x,n}, s_{y,m})$. The inner product $tau_{mn} \otimes f$ thus represents the phase at each point in the fk grid.

The power spectral density in the fk beam $V_{mn}(f)$ can then be used to calculate the fk F-statistic, similar to how the power in the time-domain stack is used to calculate $F(t)$ in Equation 3.8:

$$\mathscr{F}_{mn}(f) = \frac{N(N-1)\Re[V_{mn} \cdot V_{mn}]}{\Re[U_j \cdot U_j] - N\Re[V_{mn} \cdot V_{mn}]}. \tag{3.10}$$

Examples of VesPy fk plots for both power and $\mathscr{F}$ are shown in Figure 3.6. Fk analysis is useful for detecting and identifying seismic phase arrivals without making prior assumptions of the velocity model or the direction of arrival.

3.3 Implementation

I have developed a package of codes in Python to apply various array methods to seismic data. Python is a programming language in increasingly common use by seismologists, with platform-independence, easy-to-read syntax and line-by-line interpretation. It also has the advantage of being highly customisable with various modules that can easily be downloaded and installed. One such module is ObsPy[2], a general seismological toolkit developed by [4]. One key function of ObsPy is its ability to read in a large range of seismic data formats (including SEGY, GSE2, and SAC) and to handle them in a generalised way with a Stream and Trace internal data structure. This code package was designed to work with ObsPy, in order to be applicable on the greatest number of datasets, regardless of the initial format the data from the array is supplied in.

Array methods do however require particular metadata to be present in the data files, most importantly those that contain information about the geometry of the array. For this reason the latitude, longitude, and elevation of each receiver in the array must be recorded. The SAC data format contains a file header capable of storing this information under the `stla`, `stlo`, and `stel` header variables. To use data formats that do not natively include this information, it must be imported from a StationXML file and added to the trace's `stats` attribute.

The package comprises five modules listed in Table 3.1. Each module contains functions grouped by general purpose, as described in Tables 3.2–3.6. The code is written in such a way

[2] http://www.obspy.org/

to allow the user to import only the functions they need for their analysis. Where functions can produce graphical outputs (such as the fk analysis and vespagram functions), versions of the same function that do not produce this output are available to allow the user to pass the results into their own plotting functions.

The stacking and statistics functions return their results as NumPy arrays, so that the results are in a more useful format for plotting or further analysis. However the results in each case are still time series, and can be converted into an ObsPy `Trace` for plotting and easier comparison with the original seismograms using the `vespy.utils.traceify()` function, as illustrated in Figure 3.12.

The package is distributed using git[3] and GitHub[4]. The package comes with an installation script, `setup.py`, that will install all the modules within the user's Python installation. It was originally written for Python 2.7, but is also compatible with Python 3. Instructions for installation and running VesPy can be found in Appendix A.

3.4 Example Use Cases

3.4.1 Core phase vespagrams

Vespagrams plot seismic amplitude or power as a function of time and either the scalar slowness or backazimuth. They are therefore useful for identifying different seismic phase arrivals from their arrival time and slowness. They can therefore be applied, for example, to studies of the deep Earth —in particular to detecting core phases, which are often difficult to identify in a seismogram.

I illustrate VesPy's vespagram function using data recorded at the HI-NET array in Japan for an event in the Hindu Kush region in northeastern Afghanistan. Traces from 320 broadband stations are included, giving the resulting vespagrams a very high slowness resolution. Vespagrams are calculated in time windows around the P, PcP and PKiKP phases.

The seismic data is highpass-filtered below 2 Hz and in each vespagram cut to a time window of 20 s around the *ak135* model-predicted arrival time for each phase. The power in the 8th-root stack, calculated in a 1 s time window is plotted as a function of stack slowness and time in Figure 3.7. I observe each arrival occurring around 3 s later than predicted, and with a slowness of around 0.004 s / km slower (in the case of P and PcP) or faster (PKiKP) than in *ak135*.

[3] http://git-scm.com
[4] https://github.com/NeilWilkins/VesPy

3.4.2 Optimising corrections at IMS seismic arrays

3.4.2.1 Slowness-azimuth station corrections

Seismic arrays can show systematic biases in the measurement of the slowness and azimuth of an incoming wavefield. [10], calculated teleseismic slowness-azimuth station corrections (SASCs) for several IMS primary seismic arrays. They ascribed the cause of these biases at certain arrays to local crustal structure or instrument issues. In particular, a sinusoidal variation in slowness residuals is a feature associated with a dipping Mohorovičić discontinuity beneath the array as in, e.g., [31]. Many array methods, such as stacking, use simple 1D velocity models to calculate the delay times required to align the relevant seismic arrivals —P, S, PcP, etc. At arrays that show significant biases in local slowness structure, we should seek to apply corrections to the slowness used in creating the stack in order to properly align the signal at the array. The better the alignment of the waveforms, the higher the increase in SNR achieved in performing the stack.

Frequency-wavenumber (fk) analysis (see Section 3.2.5) provides a computationally quick method to estimate the entire 2D slowness vector (i.e. horizontal slowness and backazimuth) of a seismic phase. I have applied the fk analysis functions in the VesPy package (see Table 3.6) to a seismic dataset retrieved from the IMS Primary Seismic Array AKASG (see Figure 3.8), near Malin, Ukraine. By comparing the slowness and azimuth observed in the fk plot to those predicted by the *ak135* velocity model ([46]), I obtain SASCs for this array.

The dataset consists of a catalogue of 800 earthquakes (see Figure 3.9) at a great circle epicentral distance of 50–80° from the centre of the array in order to have signals from which a good teleseismic P phase could be recorded. I selected events with m_b 4.0–6.0. The waveforms were filtered in the pass band from 0.5 to 5.0 Hz using a 2-pole, 2-pass Butterworth filter to record teleseismic body wave arrivals.

To calculate SASCs, I performed fk analysis in a 20 s time window centred on the model-predicted arrival time for the P phase. I calculated the F-statistic for each event within a slowness grid ranging between -0.1 and $0.1\,\mathrm{s/km}$ in the north-south and east-west directions, with the location of the maximum value of F indicating the most likely slowness vector of the incoming wavefield. An example fk plot for one of the events is shown in Figure 3.10. The residuals between the observed and predicted slowness and azimuth are then used to calculate averaged SASCs for the array.

The events are binned by model-predicted backazimuth and slowness, with separate bins for each 10° backazimuth and $0.01\,\mathrm{s/km}$ slowness increment. The median backazimuth and slowness residual across all events within each bin is calculated. SASCs are only calculated for bins where there are more than five events.

To estimate the error in the SASCs I perform a bootstrapping analysis. For each array, I select one event per slowness-azimuth bin and use a random selection of stations within the array to construct a randomly-sampled virtual array. 1000 such virtual arrays are used to repeat the fk analysis, and I examine the mean and error in the mean of these results to characterise the

slowness resolution of the array in that bin. Similarly, I use the circular mean and its error to characterise the azimuthal resolution within each bin.

The SASCs calculated at AKASG are plotted as mislocation vectors in Figure 3.11. For each slowness-azimuth bin, the vector points from the centre of that bin to the slowness and azimuth to which events occurring in that bin are shifted to make the SASC. The slowness corrections at AKASG show a consistent trend across all backazimuths, where the slowness of the P-wave arrivals is around 2 s/° lower than predicted.

The improvement in the SNR achieved by stacking the traces after the corrections have been applied can be demonstrated both in the stacked waveforms themselves (Figure 3.12) and using the F-statistic. Figure 3.13 shows the F-statistic as a function of time for a series of different events recorded at AKASG, both with and without the SASCs having been applied. The mean percentage increase in $F(t)$ in the 10 s around the predicted P arrival for the corrected stacks is 197% when compared with the uncorrected traces.

I demonstrate the effect of the corrections on body wave magnitude estimates, and thus the m_b/M_S discriminant by calculating m_b for the corrected and uncorrected stacks for an earthquake on May 21st 2016 in central Alaska. In the array beam before the SASC is applied, I measure $m_b = 4.5$, and using the IDC estimate of $M_S = 3.4$, obtain an $m_b/M_S = 1.34$. With the SASC applied, Imeasure $m_b = 4.7$ and thus $m_b/M_S = 1.39$.

3.5 Summary

I present VesPy, a Python module that includes functions to perform common seismic array analysis procedures, including stacking, vespagrams and fk analysis. The code includes documentation and is available for use in any commercial or non-commercial applications.

I illustrate some example use cases of VesPy. The first shows vespagram plots for seismic core phases useful in studies of the deep Earth. The second shows an example application of VesPy's fk analysis functions in calculating slowness-azimuth corrections at a number of IMS seismic arrays, and in analysing the improvement to the signal-to-noise ratio achieved using these corrections by plotting the stacked waveforms and the F-statistic.

The software and source code is freely available via GitHub, and there is scope for the development of other array methods and related functionality.

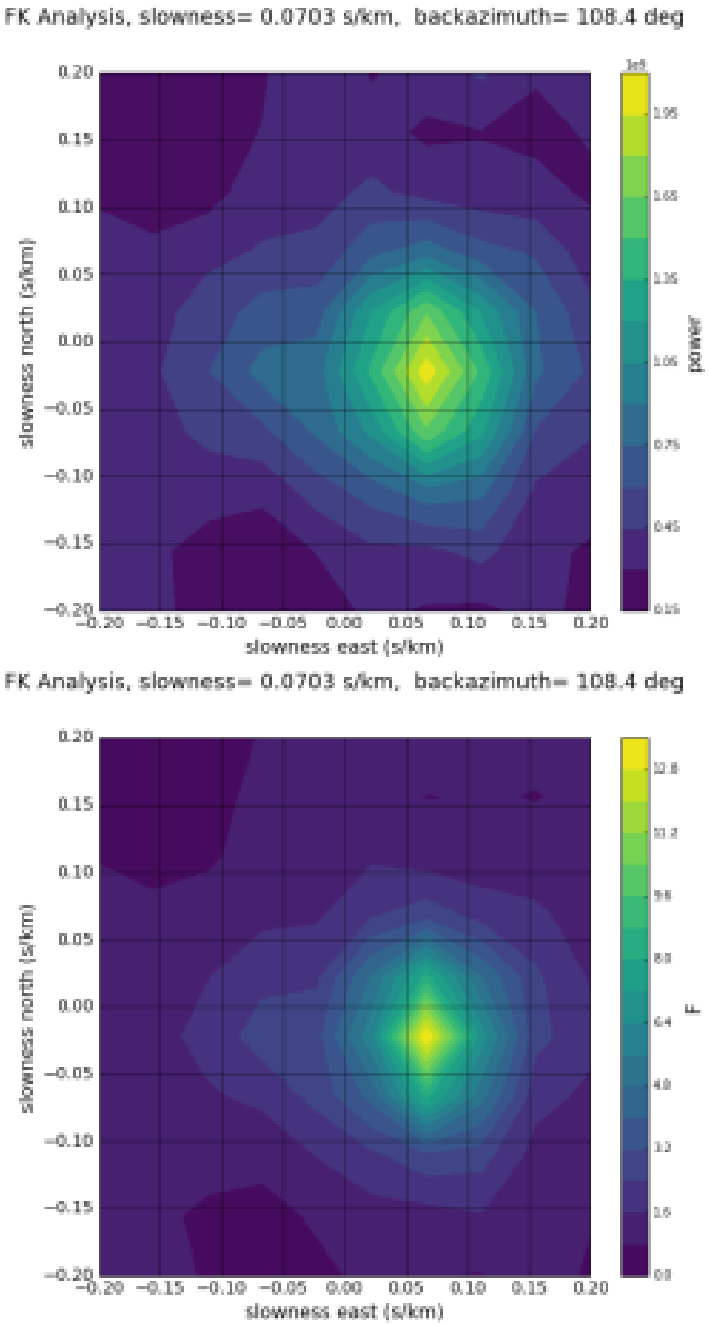

Figure 3.6: VesPy fk plots for the 2009 DPRK nuclear test recorded at SONM. Top shows power, bottom shows F-statistic. Both plots give a good slowness resolution in this case, and the slowness vector with the maximum observed power or F-statistic is the same.

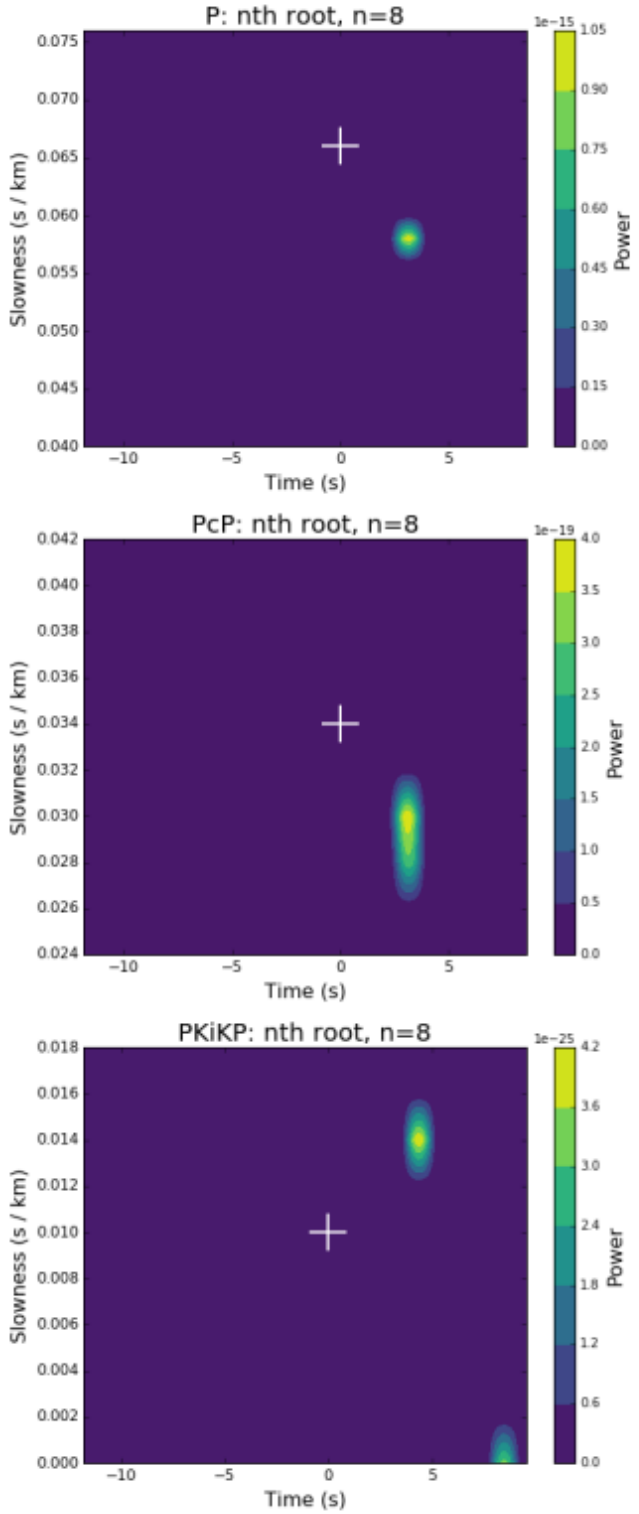

Figure 3.7: Vespagrams for a m_b=6.4 earthquake in the Hindu Kush region of Afghanistan. Each plot shows the power in the 8th-root stack as a function of slowness. Phase arrivals for the P, PcP and PKiKP phases are recorded, arriving 3 s later than predicted by *iasp91* (white crosses), and with a slowness residual of ±4 ms/°

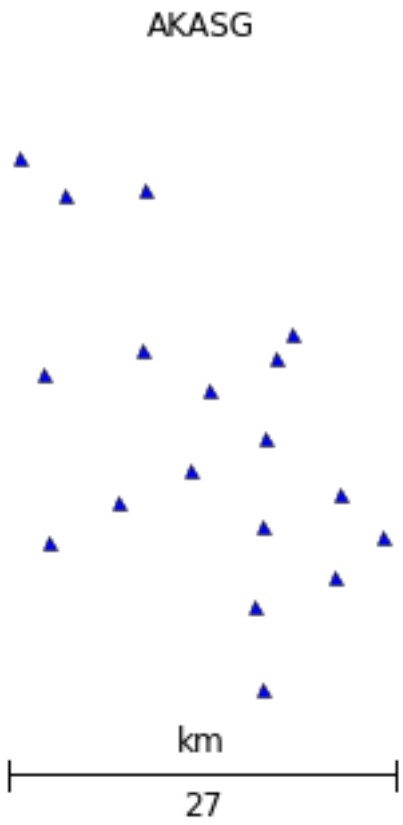

Figure 3.8: AKASG array geometry

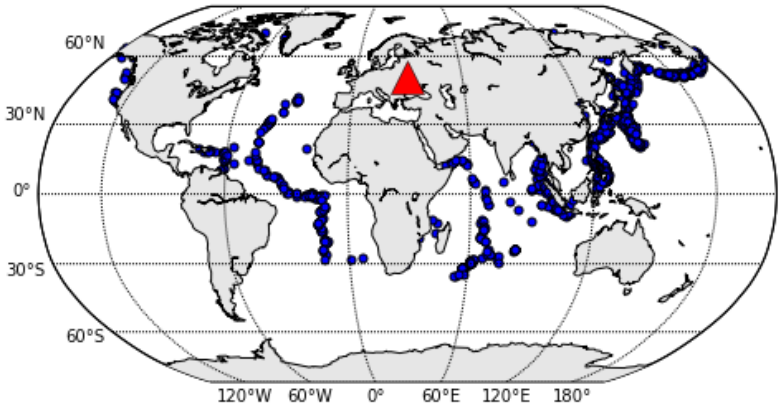

Figure 3.9: Event catalogue used as a training dataset for the slowness-azimuth corrections for the AKASG array. The catalogue includes 800 events at an epicentral distance of 50–80° from the array, with m_b 4.0–6.0.

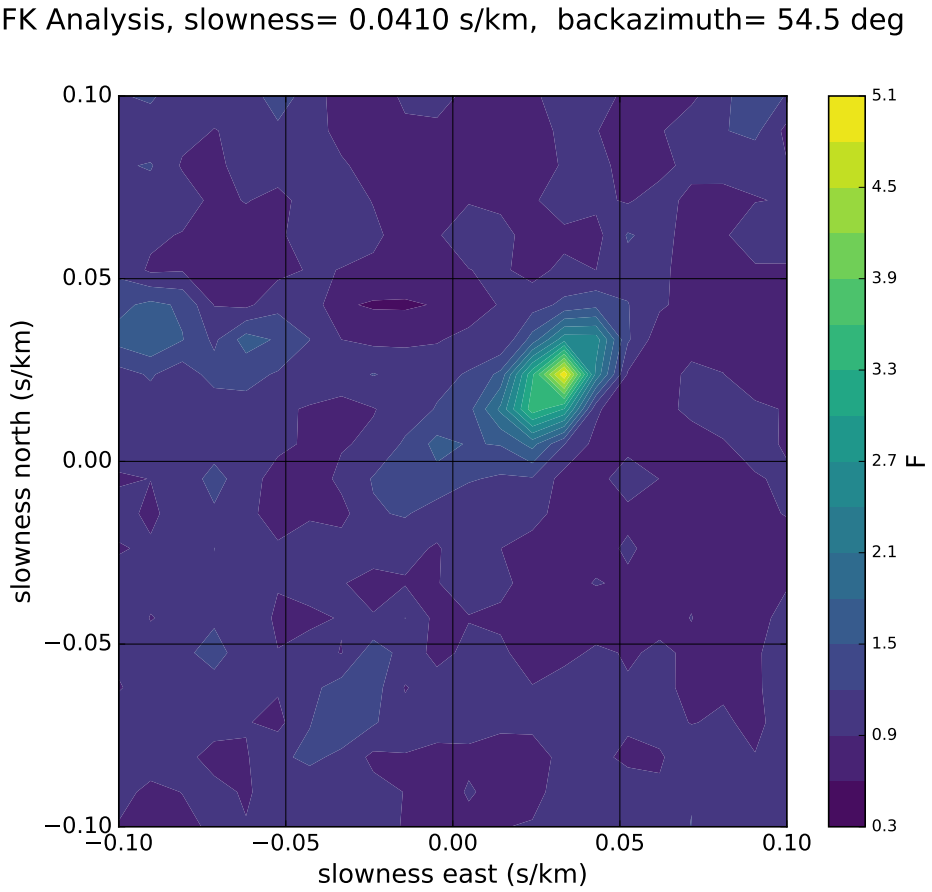

Figure 3.10: Frequency-wavenumber plot for an event located in South Korea and seismic traces recorded at AKASG. The event's magnitude is $m_b = 4.7$, and is at a great circle distance of 69° and depth of 10 km. The *ak135* predicted slowness is 0.056 s/km and backazimuth is 58.7°. The seismic data have been filtered between 0.5 and 5.0 Hz and the fk analysis conducted on a 10 s window around the predicted P arrival. The observed arrival is thus 0.015 s/km faster than predicted.

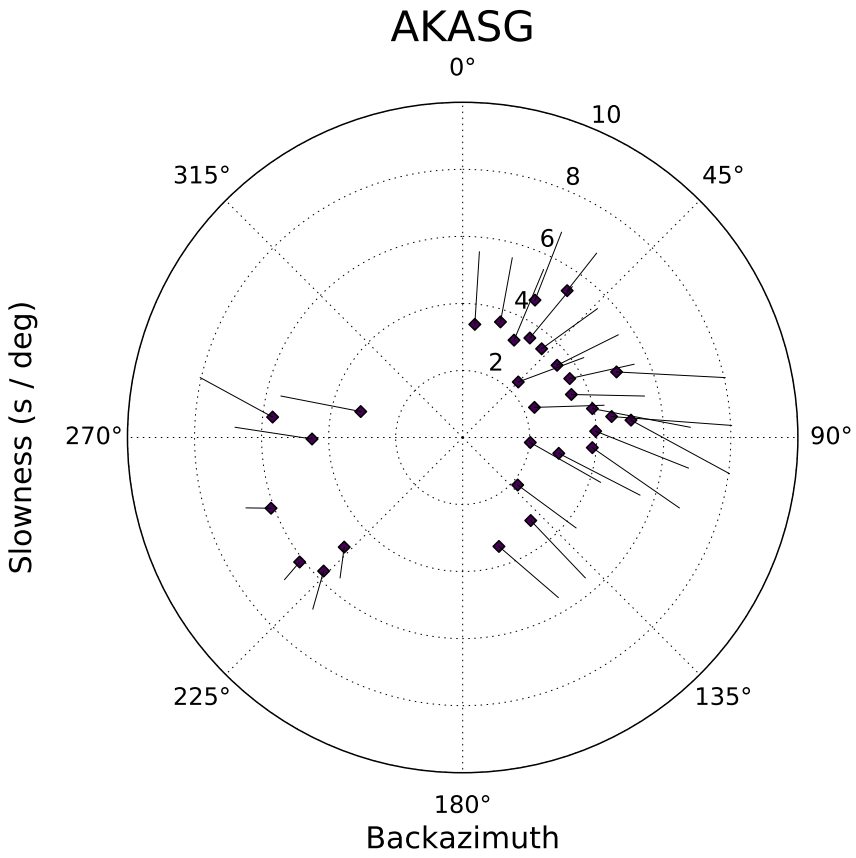

Figure 3.11: Mislocation vectors for AKASG. Vectors point from the model-predicted slowness and backazimuth to the corrected slowness and backazimuth. To apply the SASC, the *a priori* slowness vector events occurring within each slowness-azimuth bin should be shifted to the head of the arrow. The observed slowness for P phases arriving at the array are significantly faster than predicted, with a consistent slowness correction of −1.5–3.0 s/degree (−0.013–0.026 s/km).

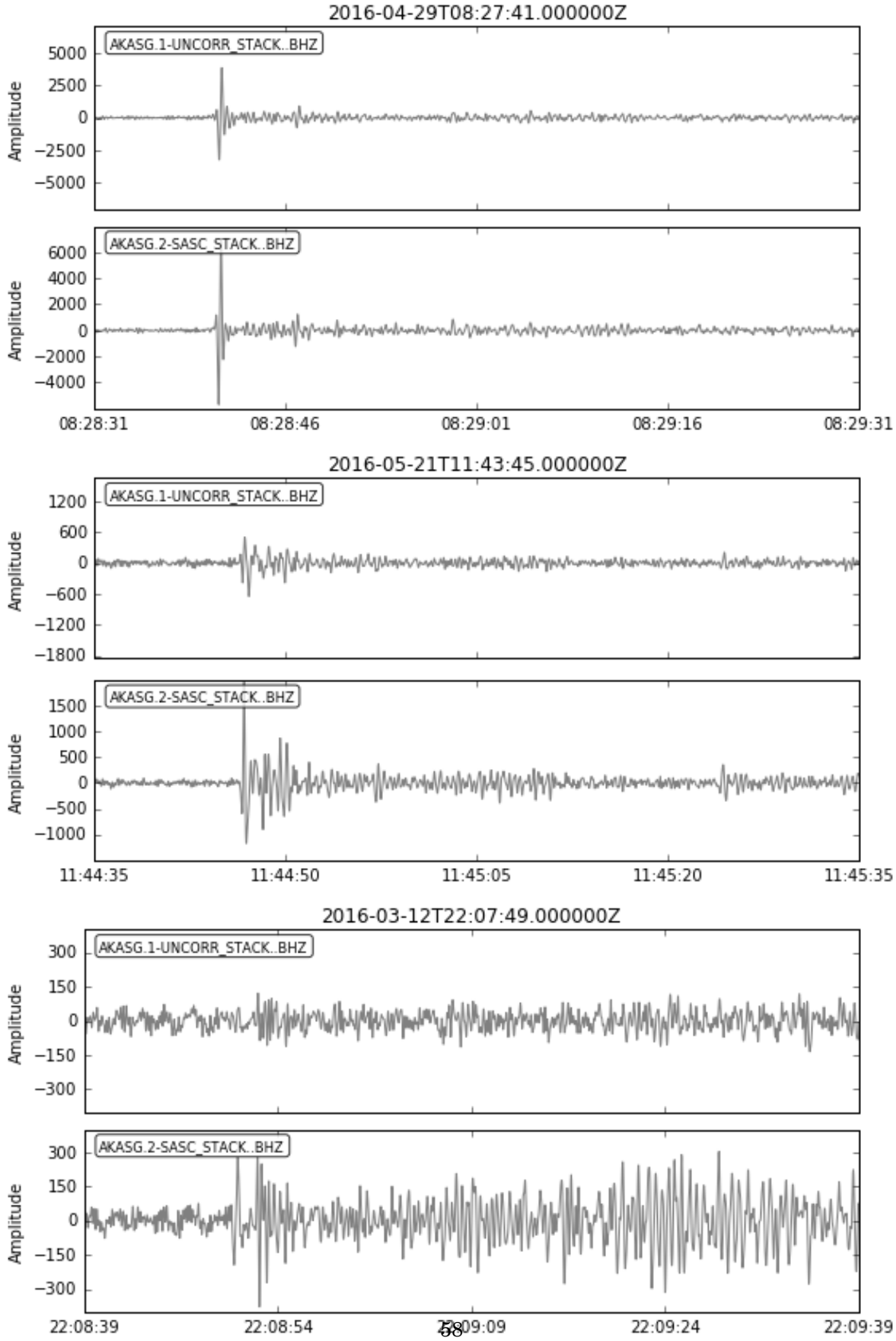

Figure 3.12: A comparison of the stacked waveforms before and after applying the SASC for three events recorded at AKASG. The signal amplitude and SNR is higher in the corrected stack.

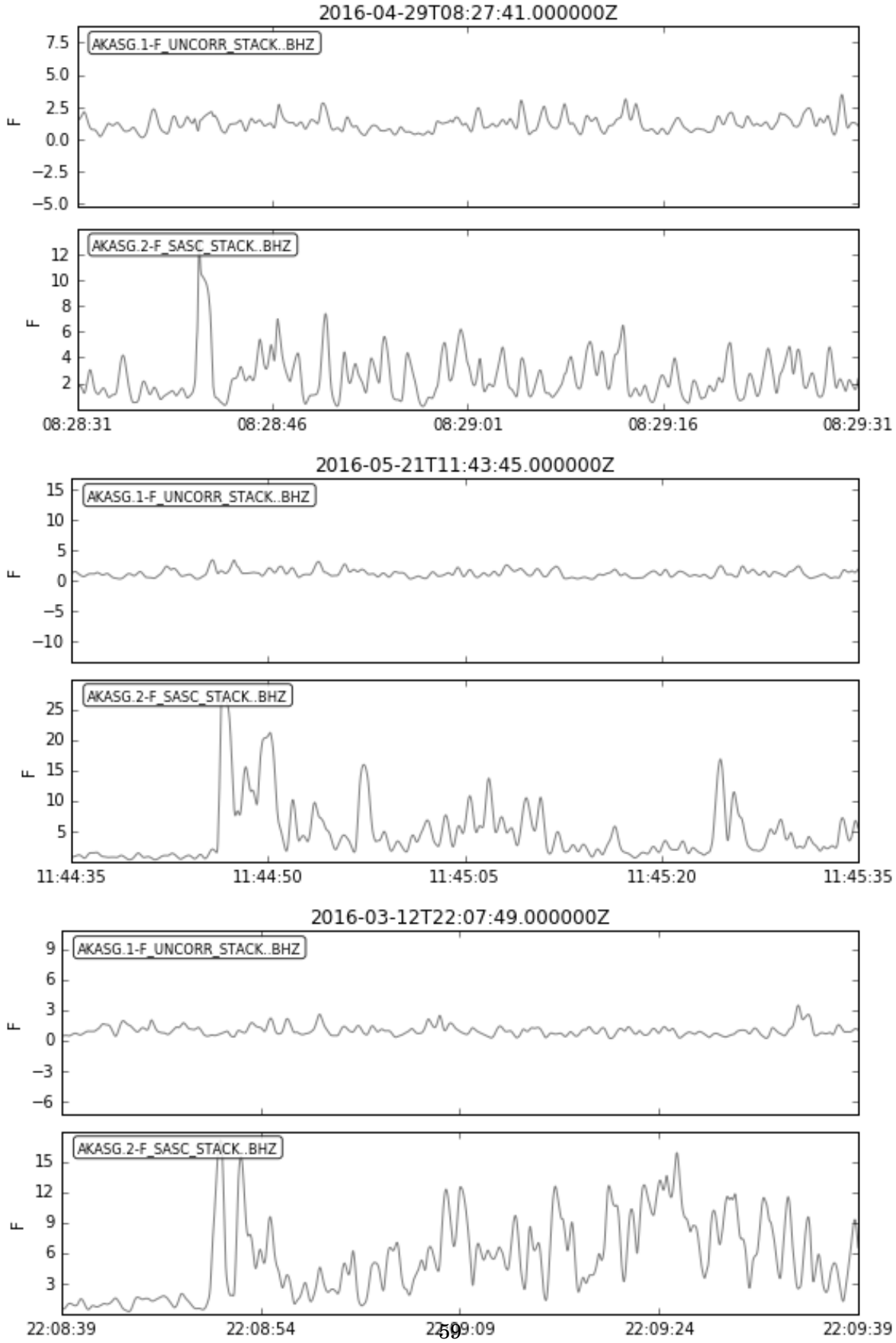

Figure 3.13: The F-statistic in the stack calculated before and after applying the SASC. F is considerably higher in the corrected stacks, further indicating the improved SNR.

Table 3.1: VesPy modules

Module name	Description
`vespy.utils`	General purpose utilities for VESPA analysis, handling array geometry, generating TauP phase arrival times, etc.
`vespy.stacking`	Functions for calculating delay times at an array, and performing various stacking methods
`vespy.stats`	Statistical functions for VESPA analysis, including semblance and F-statistic
`vespy.vespagram`	Functions for calculating and plotting vespagrams
`vespy.fk`	Functions for performing and plotting frequency-wavenumber analysis

Table 3.2: Functions in module `vespy.utils`

Function name	Description
`traceify`	Converts a NumPy array (e.g. with stacked data) into an ObsPy Trace object, using the times and header variables of a reference trace
`get_station_coordinates`	Calculates the x, y, z coordinates of stations in a seismic array relative to a reference point for a given stream of SAC seismographic data files.
`plot_array_map`	Produces a plot of the seismic array geometry for the stations present in the stream
`get_arrivals`	Returns complete arrival information (name, travel time, slowness) from theoretical velocity model (default *ak135*) for the event and stations contained in the stream
`get_first_arrival`	Returns arrival information (name, travel time and slowness) for only the first arrival predicted by the velocity model (default *ak135*)

3.6 Tables of functions in VesPy

The following tables describe the various functions included in the VesPy code package. The functions are divided into modules based on application, as described in Table 3.1.

Table 3.3: Functions in module `vespy.stacking`. All stack functions return a NumPy array.

Function name	Description
`get_shifts`	Calculates delay time shifts (in integer number of data points) based on array geometry and slowness and backazimuth of incoming wavefield
`linear_stack`	Computes the linear stack for the stream for the given slowness and backazimuth
`nth_root_stack`	Computes the nth-root stack for the stream for given n, slowness and backazimuth
`phase_weighted_stack`	Computes the phase-weighted root stack for the stream for given sharpness, slowness and backazimuth

Table 3.4: Functions in module `vespy.stats`

Function name	Description
`semblance`	Calculates the semblance in the linear stack of the stream for a given moving time window
`f_vespa`	Calculates the F-statistic in the linear stack of the stream for a given moving time window
`power_vespa`	Calculates the power in the linear stack of the stream for a given moving time window
`n_power_vespa`	Calculates the power in the nth-root stack of the stream for a given moving time window
`f_stat`	Calculates the F-statistic in a provided stack for the stream for a given moving time window

Table 3.5: Functions in module `vespy.vespagram`

Function name	Description
`vespagram`	Calculates the vespagram for the stream over a given slowness range, for a single backazimuth, using the statistic specified (default power, or amplitude, semblance, F)
`plot_vespagram`	Plots the vespagram for the same parameters as the `vespagram` function
`f_vespagram_theoretical_arrivals`	Plots a vespagram of the F-statistic, with markers plotting the *ak135* predicted arrivals
`vespagram_backazimuth`	Calculates the vespagram for the stream over a given backazimuth range, for a single horizontal slowness, using the statistic specified (default power, or amplitude, semblance, F)
`plot_vespagram_backazimuth`	Plots the vespagram for the same parameters as the `vespagram_backazimuth` function
`plot_stack_vespagram`	Plots the slowness stacked waveforms as a vespagram for the stream over a given slowness range

Table 3.6: Functions in module `vespy.fk`

Function name	Description
`fk_analysis`	Performs frequency-wavenumber analysis on the stream of N traces in the given frequency and time windows. Returns a NumPy array with size $N \times N$ giving the chosen statistic (default: power) for each point on the slowness grid.
`fk_slowness_vector`	Performs frequency-wavenumber analysis, returning the slowness and backazimuth of maximum power
`fk_plot`	Performs frequency-wavenumber analysis and produces a plot showing the chosen statistic (default: power) as a function of s_y, s_x

3.7 Code excerpt

The following is an excerpt from the code in the package VesPy, from the `vespy.vespagram` module, including the function to calculate the vespagram. The full code is freely available online at https://github.com/NeilWilkins/VesPy.

3.7.1 Vespagram function

```
# vespagram.py
# VesPy Vespagram Module
# vespy.vespagram

import numpy as np
import matplotlib.pyplot as plt
import matplotlib.cm as cm
from obspy.taup import TauPyModel
from obspy.geodetics import locations2degrees

from vespy.utils import G_KM_DEG
from vespy.stacking import linear_stack, nth_root_stack, phase_weighted_stack
from vespy.stats import n_power_vespa, f_vespa, pw_power_vespa

def vespagram(st, smin, smax, ssteps, baz, winlen, stat='power',
    phase_weighting=False, n=1):
    '''
    Calculates the vespagram for a seismic array over a given slowness range,
        for a single backazimuth, using the statistic specified.

    The chosen statistic is calculated as a function of time (in s) and
        slowness (in s/km). This may be:-

    * 'amplitude' - the raw amplitude of the linear or nth-root stack at each
        time and slowness step;
    * 'power' - the power in the linear or nth-root beam calculated over a time
        window (length winlen) around each time step for each slowness;
    * 'F' - the F-statistic of the beam calculated over a time window (length
        winlen) around each time step for each slowness.

    Parameters
    ----------
    st : ObsPy Stream object
        Stream of SAC format seismograms for the seismic array, length K = no.
            of stations in array
```

```
    smin : float
        Minimum magnitude of slowness vector, in s / km
    smax : float
        Maximum magnitude of slowness vector, in s / km
    ssteps : int
        Integer number of steps between smin and smax for which to calculate
            the vespagram
    baz : float
        Backazimuth of slowness vector, (i.e. angle from North back to
            epicentre of event)
    winlen : int
        Length of Hann window over which to calculate the power.
    stat : string
        Statistic to use for plotting the vespagram, either 'amplitude',
            'power', or 'F'

    Returns
    -------
    vespagram_data : NumPy array
        Array of values for the chosen statistic at each slowness and time
            step. Dimensions: ssteps*len(tr) for traces tr in st.
    '''

    assert stat == 'amplitude' or stat == 'power' or stat == 'F', "'stat'
        argument must be one of 'amplitude', 'power' or 'F'"

    vespagram_data = np.array([])

    try:
        if stat == 'amplitude':
            if phase_weighting:
                vespagram_data = np.array([phase_weighted_stack(st, s, baz, n)
                    for s in np.linspace(smin, smax, ssteps)])
            else:
                vespagram_data = np.array([nth_root_stack(st, s, baz, n) for s
                    in np.linspace(smin, smax, ssteps)])

        elif stat == 'power':
            if phase_weighting:
                vespagram_data = np.array([pw_power_vespa(st, s, baz, n,
                    winlen) for s in np.linspace(smin, smax, ssteps)])
            else:
```

```
                vespagram_data = np.array([n_power_vespa(st, s, baz, n, winlen)
                    for s in np.linspace(smin, smax, ssteps)])

        elif stat == 'F':
            vespagram_data = np.array([f_vespa(st, s, baz, winlen, n) for s in
                np.linspace(smin, smax, ssteps)])

    except AssertionError as err:
        raise err

    return vespagram_data
```

CHAPTER 4

NORTH KOREAN NUCLEAR TESTS

Since the first announced test in North Korea in October 2006, nuclear weapons testing has once again been brought to prominence, leading to heightened political tension in East Asia and internationally. A total of six announced underground nuclear explosions and over 100 ballistic missile tests [57] have been conducted in the country to date. The only nuclear tests to have been detected in the last twenty years, the North Korean tests provide an important test of the ability of the International Monitoring System (IMS) to reliably detect and verify underground nuclear testing.

In this chapter, I use data from earthquakes and underground tests around the Korean peninsula to develop slowness corrections for IMS seismic arrays to improve the detection and verification of seismic events in the region surrounding the North Korean nuclear test site.

4.1 Introduction

The primary seismic arrays of the International Monitoring System (IMS) have been important sources of data for the detection and verification of the six announced nuclear tests conducted by the Democratic People's Republic of Korea (DPRK) from 2006–17 ([2], [85]). The DPRK tests, listed in Table 4.1, have been the only announced nuclear explosions conducted since the IMS has been operational, providing an important test for the system's capabilities. The seismic events associated with the six tests have ranged in magnitude (and thus the derived explosive yields of the tests themselves) from 4.1 for the first test in October 2006 to 6.1 for the most recent test in September 2017 [56]. Accurate magnitudes for suspected or announced nuclear tests are important in estimating the probable yield of the nuclear device, and for discriminating underground explosions from naturally-occurring earthquakes using m_b/M_S.

Table 4.1: Nuclear tests conducted in North Korea to date

Event	Date and time (UTC)	m_b (CTBTO)	Estimated Yield (NORSAR)[56]
2006	2006-10-09 01:35:27	4.0	~0.5 kT
2009	2009-05-25 00:54:43	4.5	~1–3 kT
2013	2013-02-12 02:57:51	4.9	10 kT
2016a	2016-01-06 01:30:01	4.9	5 kT
2016b	2016-09-09 00:30:01	5.1	10–15 kT
2017	2017-09-03 03:30:01	6.1	200–300 kT

In addition to the six announced tests, there is some evidence that a small nuclear explosion was carried out in the DPRK in May 2010. This claim was based on radionuclide evidence by de Geer [23]. No seismic event was associated with the radionuclide data until Zhang and Wen [84] reported evidence of a low-magnitude (m_b 2.9) seismic event, identified as an explosion, recorded at Chinese seismic arrays within 200 km of the test site. The identification of this event with the radionuclide data is a matter of ongoing debate (e.g. [49], [24]), and the CTBTO has not formally confirmed that a 2010 nuclear test took place. However, the uncertainty illustrates the importance of monitoring for suspicious seismic activity even at low magnitudes.

While events of the magnitude of the largest test are readily detectable by IMS seismic stations, it is still necessary for the system to be able to detect smaller magnitude events in order to ensure a robust verification regime for the Comprehensive Test-Ban Treaty Organisation (CTBTO). To assist this, beamforming can be performed at seismic arrays, with the stacked seismic waveforms showing an improved signal-to-noise ratio (SNR) when compared with the individual traces recorded at each station [66]. Stacking techniques require a prior assumption of the velocity model for the seismic phases; at the CTBTO's International Data Centre (IDC), the IASPEI-1991 (*iasp91*) one-dimensional velocity model [45] is used in compiling their event bulletins. However, several IMS primary seismic arrays show significant deviations in the observed arrival times of seismic phases from this model. To account for this, slowness-azimuth station corrections (SASCs) have been developed for many of the arrays in the network [10]. These corrections are applied for events based on their distance (the *a priori* slowness of the seismic phase) and azimuth from the event's origin. By stacking the traces using the corrected slowness, the phases align more closely between each receiver in the array, and the SNR enhancement of the stack is improved.

Stacked waveforms can also be used to make estimates of the seismic magnitude of suspected events. Accurate magnitude estimates are important both for discriminating between explosions and earthquakes using the m_b/M_S ratio and in estimating the explosive yield of a nuclear test e.g., [16].

In this chapter, I develop SASCs for four IMS primary arrays at regional distances to the DPRK nuclear test site at Punggye-ri, the recorded location of all six announced nuclear tests conducted in the country to date [33]. In doing so I characterise the velocity structure of the

region, and show the enhancement of the SNR of the resulting stacks, and the improvement to the measurement of seismic discriminants this produces.

4.2 Background

4.2.0.1 Slowness-azimuth station corrections

Seismic waves from a distant source propagate across an array approximately as planar wavefronts with a horizontal velocity $\vec{u}$. This velocity can also be expressed as a scalar wavespeed u and a backazimuth β defined as the angle measured clockwise from North back to the source (in contrast with the azimuth α, which is defined from the source to the receiver).

When describing a particular seismic phase, it is useful to define the ray parameter, or slowness, of a phase, which remains constant along the entire raypath. The slowness is given by

$$s = u \sin\theta, \tag{4.1}$$

where θ is the incidence angle of the ray from the vertical. Thus s depends only on the type of seismic wave (e.g., P- or S-waves), and the raypath, which depends on the depth of the source and the distance to the array within the chosen velocity model. For a given seismic phase originating from a certain location, s and β are therefore constant. For example, the first P-wave arrivals generated by any number of underground explosions at a certain test site would be expected to arrive at the array with the same specific slowness and backazimuth. The slowness of the phase arrivals can be predicted using a velocity model.

However, there is often a difference between the predicted slowness of those arrivals and the slowness that is observed. Slowness-azimuth station corrections (SASCs) are designed to compensate for this difference when performing waveform stacking.

After Bondár [10], slowness-azimuth station corrections (SASCs) consist of a correction to the slowness and a correction to the backazimuth as follows:

$$s_{corrected} = s_{observed} - s_{correction}, \tag{4.2}$$

$$\beta_{corrected} = \beta_{observed} - \beta_{correction}. \tag{4.3}$$

SASCs can be illustrated using a mislocation vector which shifts the slowness vector predicted by the *a priori* velocity model onto the slowness vector actually observed at the array for the targeted phase. The SASCs measured across all azimuths and slownesses (corresponding with the great circle distance to the seismic source) can often show a trend, characterising the velocity structure beneath the array. The observed deviations from the global velocity model can indicate

Table 4.2: IMS primary seismic arrays regional to the Punggye-ri test site

Array	Location	No. stations	Aperture	Distance from test site (degrees)
KSRS (PS31)	Wonju, South Korea	17	13 km	3.9 (437 km)
MJAR (PS22)	Matsushiro, Japan	11	12 km	8.5 (950 km)
SONM (PS25)	Songino, Mongolia	10	7 km	17.9 (1936 km)
USRK (PS37)	Ussuriyisk, Russia	9	5 km	3.6 (402 km)

the geological structure of the array site, for example a sinusoidal variation has been shown to be evidence of a dipping Moho beneath the array [31]. Extreme differences in topography between stations can also have an influence, although this is generally a problem for large seismic networks used as arrays, such as the Kyrgyz Broadband Seismic Network [42]. The deviations can also be due to instrumental faults, and can indicate a systematic bias caused by a misalignment, miscalibration, or a lack of synchronisation of the seismometers deployed at certain stations within the array [10].

Slowness corrections are important to consider in seismic array analysis. The empirical slowness corrections calculated by Gibbons *et al.* [33] improved relative location estimates for the DPRK nuclear tests. Correctly applying SASCs can improve phase detection [10], and affect the results of array procedures that rely on accurate knowledge of the backazimuth [42].

4.3 Data and Methods

I have chosen four IMS primary seismic arrays at regional distances (i.e. less than 2000 km) from the Punggye-ri Nuclear Test Site in the DPRK, pictured in Figure 4.1. The arrays are situated in South Korea, Russia, Japan and Mongolia, and their details are in Table 4.2. The arrays have a range of aperture sizes, with SONM and USRK being small, circular arrays and MJAR and KSAR being larger, with irregular deployments of their elements. I choose the IMS arrays, as they contribute directly to the IDC's assessment of suspicious events. These regional arrays, being close to the test site, allow the regional velocity structure to be investigated in more detail than teleseismic stations can.

As a training catalogue for calculating the SASCs I use the International Seismology Centre's (ISC) Reviewed Event Bulletin (REB) [41], including all the seismic events with $m_b > 3.0$ (to encompass the expected magnitude range from low yield, or decoupled nuclear tests as well as larger explosions) recorded with locations within 2° of the test site. The catalogue spans the period from 1st January 2005 to 30th June 2017, and includes 393 total events, including those from the announced nuclear tests conducted during this period. SONM and MJAR were operational in 2005, and recorded 391 and 386 of the events, respectively. Data from KSRS was available starting in November 2006, with 343 of the events recorded. The USRK array was not operational until December 2008, recording 271 of the events. KSRS and USRK therefore do not

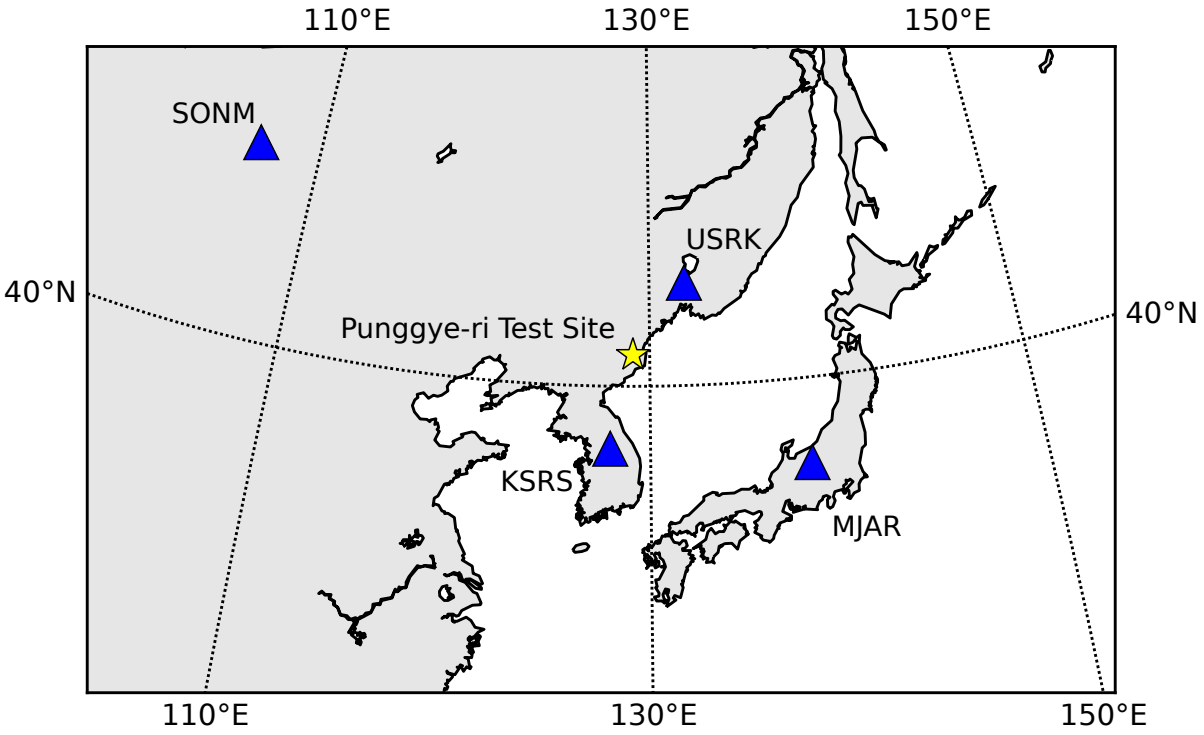

Figure 4.1: IMS primary seismic arrays within 2000 km of the North Korean Nuclear Test Site at Punggye-ri.

have waveform data for the 2006 announced nuclear test in the DPRK. The full event catalogue is plotted in Figure 4.3.

For each event, I use the TauP raypathing program [20], implemented in ObsPy, to calculate the expected arrival time at each array in the *iasp91* one-dimensional velocity model, which is used by the IDC for their event bulletins. I use the locations of the events, as recorded in the IDC REB, to calculate the backazimuth from each station in the array back to the event source. I then obtain seismic data from the arrays where the event is recorded, and cut the traces to a 10 s window around the predicted arrival of the P phase. The data is bandpass-filtered between 0.5 and 5.0 Hz, the expected frequency range for body wave arrivals at this distance, above the oceanic microseismic noise.

To measure the slowness vector of the observed P-wave arrivals, I perform fk analysis [14] on the data. I calculate the fast Fourier transform of the traces, and stack them in the frequency-wavenumber domain on a slowness grid of $N \times N$ points, where N is the number of stations in the array. At each point of the slowness grid, I calculate the F-statistic [7] in the fk stack (see Equation 3.8). The point on the grid (s_x, s_y) where F is maximised in the time window and frequency band around the expected P-wave arrival is used as an indication of the best beam, and a measure of the slowness vector of the observed arrival. This procedure is repeated for every event in the catalogue, at each of the four arrays (where the data is available). Figures 4.4 and 4.5 show fk diagrams (Equation 3.10) for two events recorded at each of the arrays. At some

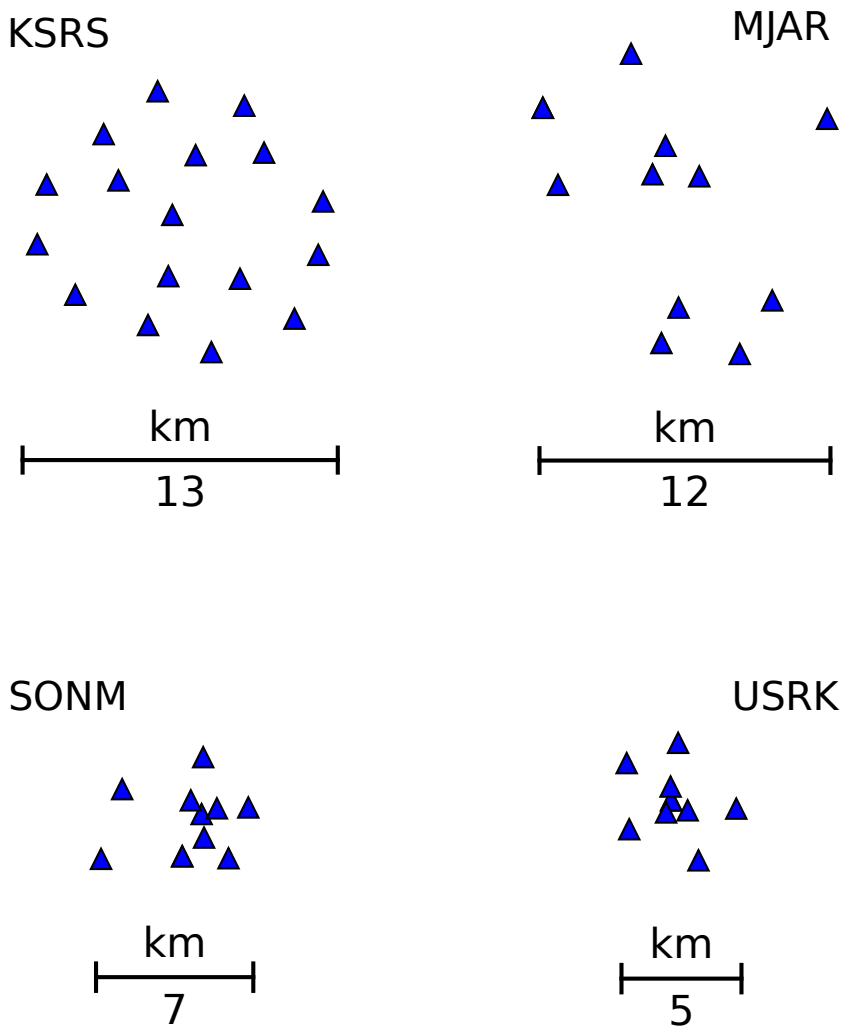

Figure 4.2: Array geometries for the arrays used in this study.

arrays the P-wave arrival is clearly recorded, with a sharp maximum in the F-statistic that is offset from the predicted arrival. In others where the SNR is low, no clear F maximum can be seen. In these cases, the event is not used when calculating the SASC.

For each event, the fk analysis gives a slowness and backazimuth based on the observed maximum F-statistic recorded at the array. The residuals between the observed and predicted slowness and backazimuth are taken for each event. These residuals indicate how far the observed arrival appears to be shifted from the predicted slowness vector.

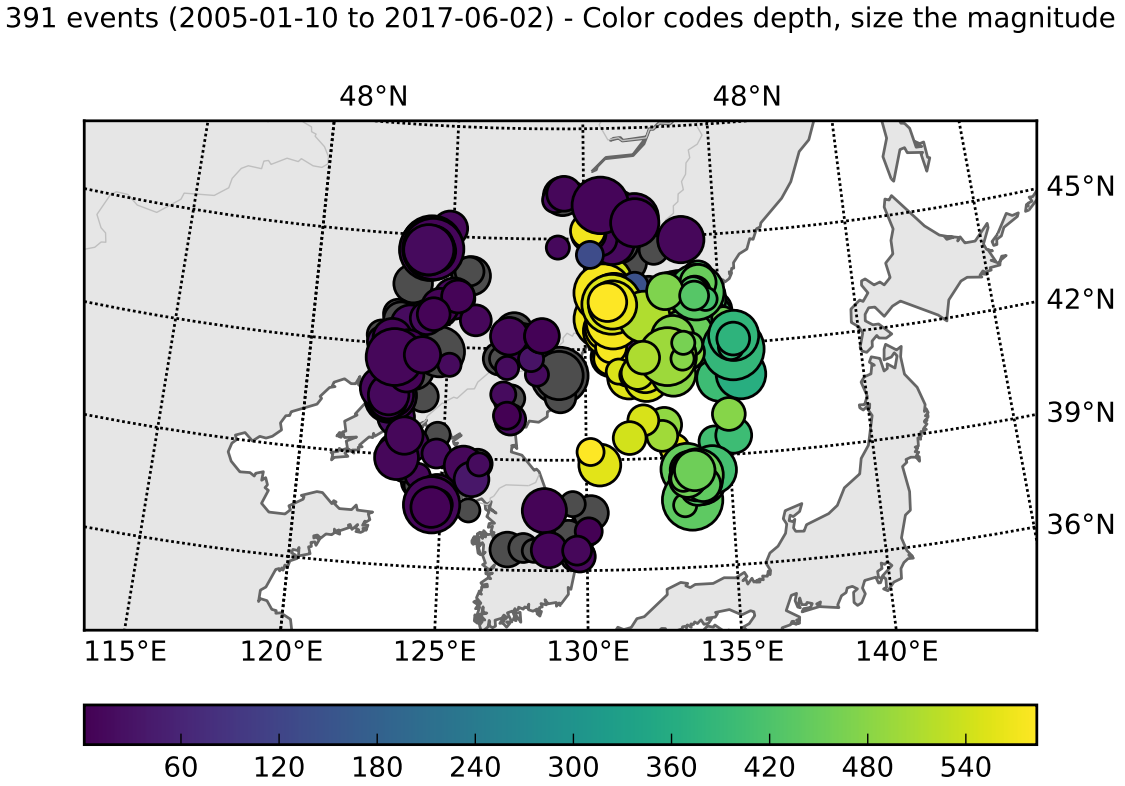

Figure 4.3: Event catalogue. All events are within 2° of the test site at Punggye-ri in the DPRK. The magnitudes range from m_b3.0–5.1. Events are colour coded based on depth.

The events are binned for each array in a two-dimensional grid of backazimuth and *a priori* slowness. The grid has dimensions of 36×20, corresponding to a backazimuthal width of 10° and a slowness width of 0.1 s/°. The fk results for events within the bin are combined, with the median slowness and azimuth residuals giving the SASC for the bin. Bins in which the total number of events recorded was less than three, or for which the backazimuth residual $\Delta\beta > 15°$ or the slowness residual $\Delta s > 5\,\mathrm{s}/°$ are considered unrealistic, and no SASC is calculated for the bin. The number of slowness bins representing the area around the test site depends on the distance of the array from the region. For the most remote array, SONM, this means 8 SASCs are obtained for the region. I obtain 17 for MJAR, which is closer to the test site. At USRK and KSRS, the region can be imaged with a greater resolution, and I obtain 38 and 40 SASCs respectively. The event count within each bin is shown in Figure 4.6.

In order to estimate the error in the SASCs calculated above, I use a bootstrapping method. For each array, I resample the seismic traces (with replacement) to assemble a virtual array with the same number of stations. I then perform the same fk analysis procedure for each array, using one event in each bin. By calculating these fk plots for 1000 such virtual arrays, and the distribution of the residuals of them all, I obtain estimates for the standard error for both the backazimuth and slowness corrections in each bin.

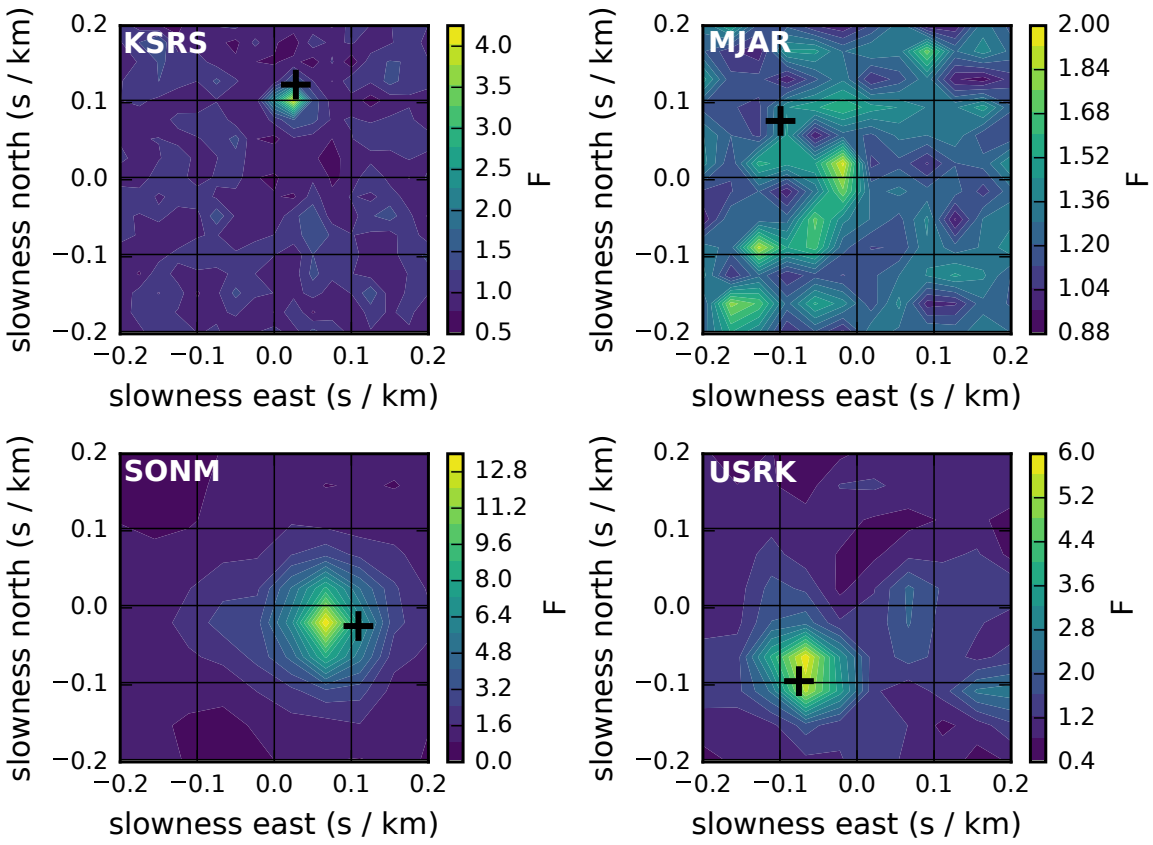

Figure 4.4: Frequency-wavenumber plots for the m_b 4.5 seismic event following the 2009 DPRK announced nuclear test, as recorded at each of the four arrays. The F-statistic is plotted as a function of slowness in the north-south and east-west directions. The maximum in F shows the observed phase arrival, while the cross represents the theoretical arrival for the event. For this event, good residuals are obtained for each array except MJAR, where the data has a low SNR.

4.4 Results

Using the data and method above, I have obtained slowness-azimuth station corrections for the four arrays, which are plotted as mislocation vectors as a function of slowness and backazimuth in Figure 4.7. For three of the arrays, these include corrections for the slowness and azimuth expected for arrivals from the Punggye-ri test site. At USRK, no correction for the test site was obtained, as the residuals obtained for the event catalogue did not pass the quality control criteria described in Section 4.3.

At SONM, the most distant array from the region, the event catalogue represents a much more narrowly defined range of incident slownesses and backazimuths than at the other arrays. Nevertheless, the SASCs obtained are consistent for the entire catalogue, with arrivals from almost all bins being observed to arrive with a slowness around 8 s/° lower than predicted by *iasp91*.

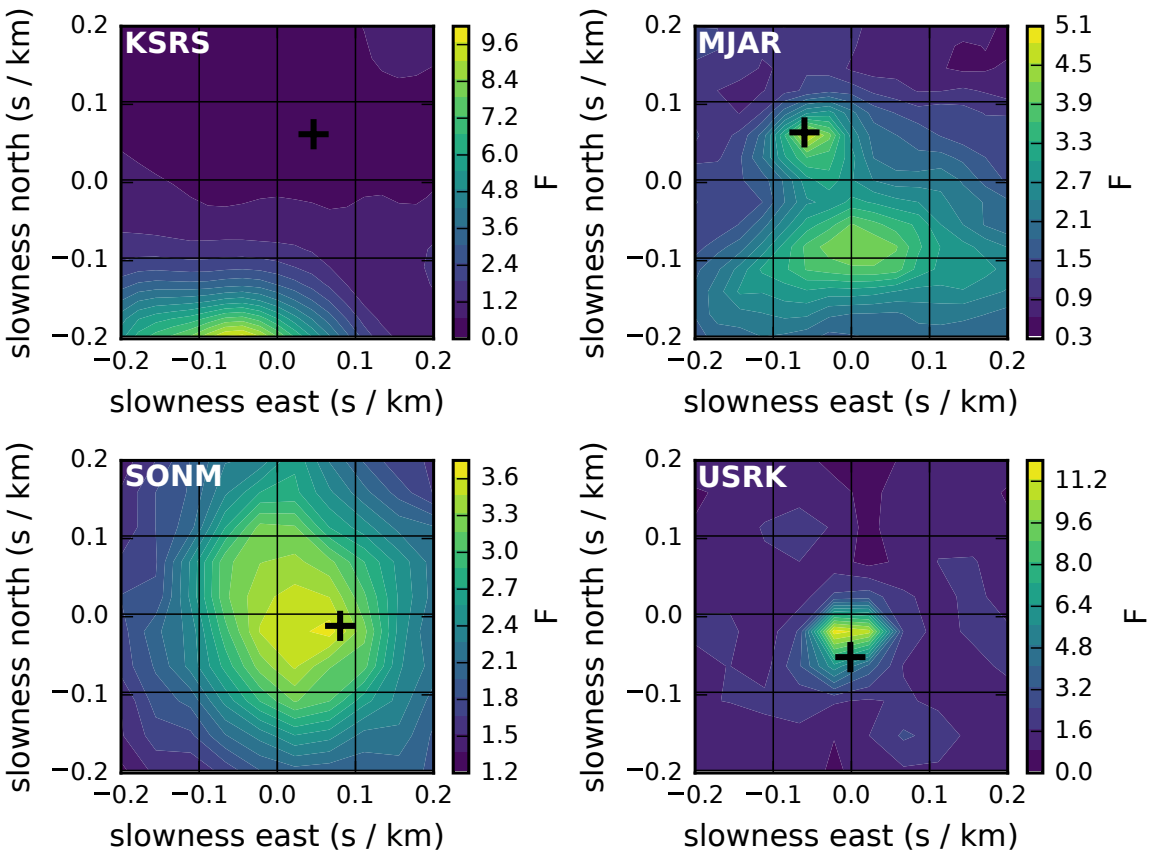

Figure 4.5: Frequency-wavenumber plots for the m_b 3.3 earthquake in the Sea of Japan recorded at each of the four arrays on the 6th August 2011. The F-statistic is plotted as a function of slowness in the north-south and east-west directions. The maximum in F shows the observed phase arrival, while the cross represents the theoretical arrival for the event. For this event, good residuals are obtained for each array except KSRS, where noise probably from a nearby microseismic event occurring in the opposite direction dominates the F-statistic.

At USRK events from different regions show different corrections. Events in the 240–290° backazimuth range, which have their origins in northeastern China, show observed arrivals around 4 s/° faster than predicted. Arrivals with backazimuths between 120–180°, representing events located in the Sea of Japan show smaller corrections of −3 s/°.

At MJAR, events (with epicentres closer to the array) with slowness greater than 10 s/° have SASCs of around +5 s/°, the only consistently positive SASCs observed at any of the four IMS arrays.

At KSRS, situated within the distribution of events used in this study, the corrections vary inconsistently depending on the slowness and backazimuth.

To measure the signal improvement that results from applying the SASCs, I compare the linear stacks calculated with the *iasp91* velocity model before and after the corrections are

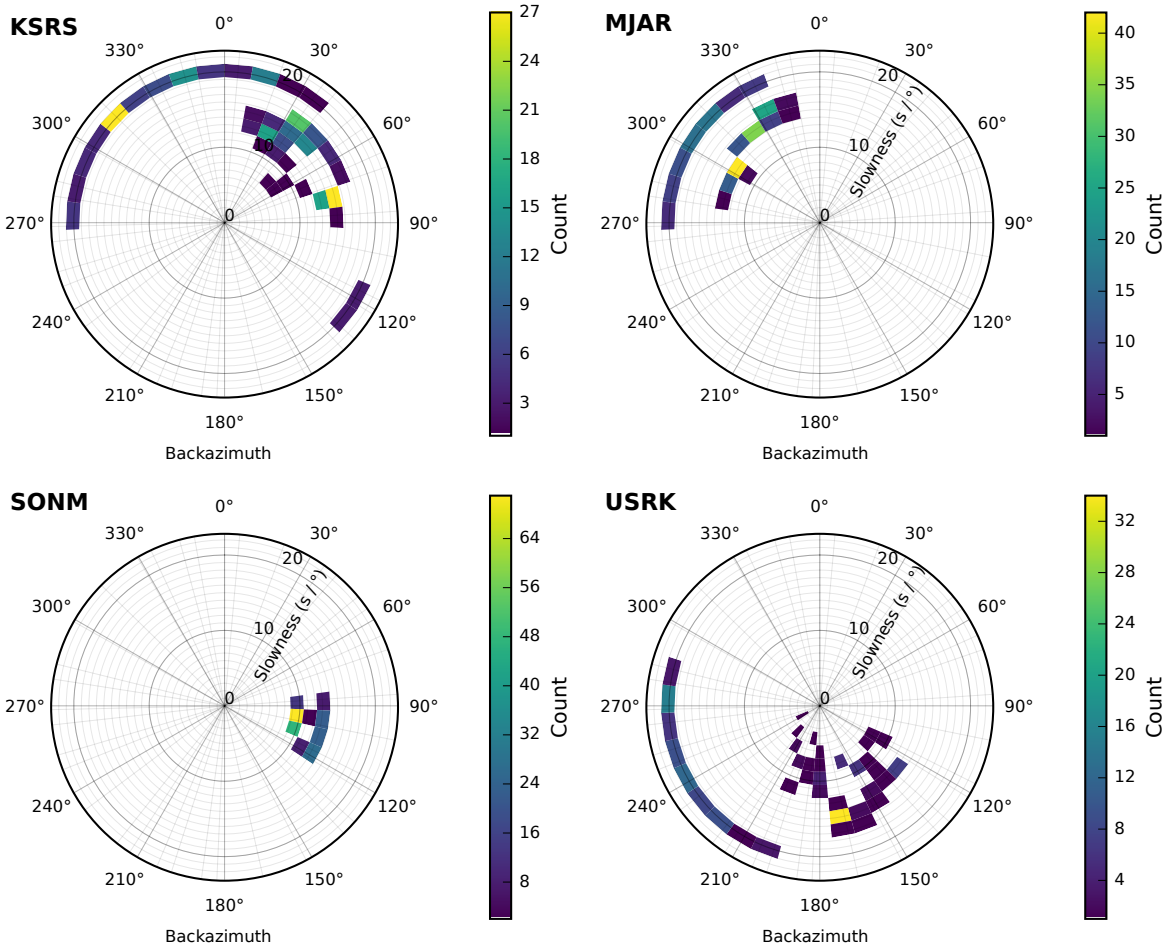

Figure 4.6: The number of events within each bin contributing to the SASCs for each array. KSRS and MJAR, as the longest-operating arrays, recorded more events in the region overall. At the more distant arrays, fewer bins are represented, but with higher event counts.

Table 4.3: Events used in F-statistic comparison. Depths not reported by the ISC for the events in China and South Korea.

Event	Date and Time (UTC)	Location	Latitude	Longitude	Depth (km)	m_b
1	2009-05-25 00:54:43	North Korea	41.3	129.0	0	4.5
2	2013-07-23 09:18:09	Sea of Japan	42.0	133.2	500	4.1
3	2016-02-22 21:03:58	South Korea	37.0	129.4	—	3.8
4	2011-12-26 13:34:09	Northeastern China	42.4	127.2	—	3.6

applied. By comparing the F-statistic in both stacks, I estimate the SNR enhancement. In each case the stack is assembled using the corrected and uncorrected delay times for the first P-wave arrival. F is calculated in a 1 s moving time window. The events used in the comparison are described in Table 4.3 and pictured in Figure 4.8.

Figures 4.9– 4.12 show the results for the DPRK announced nuclear test in 2009. No correction

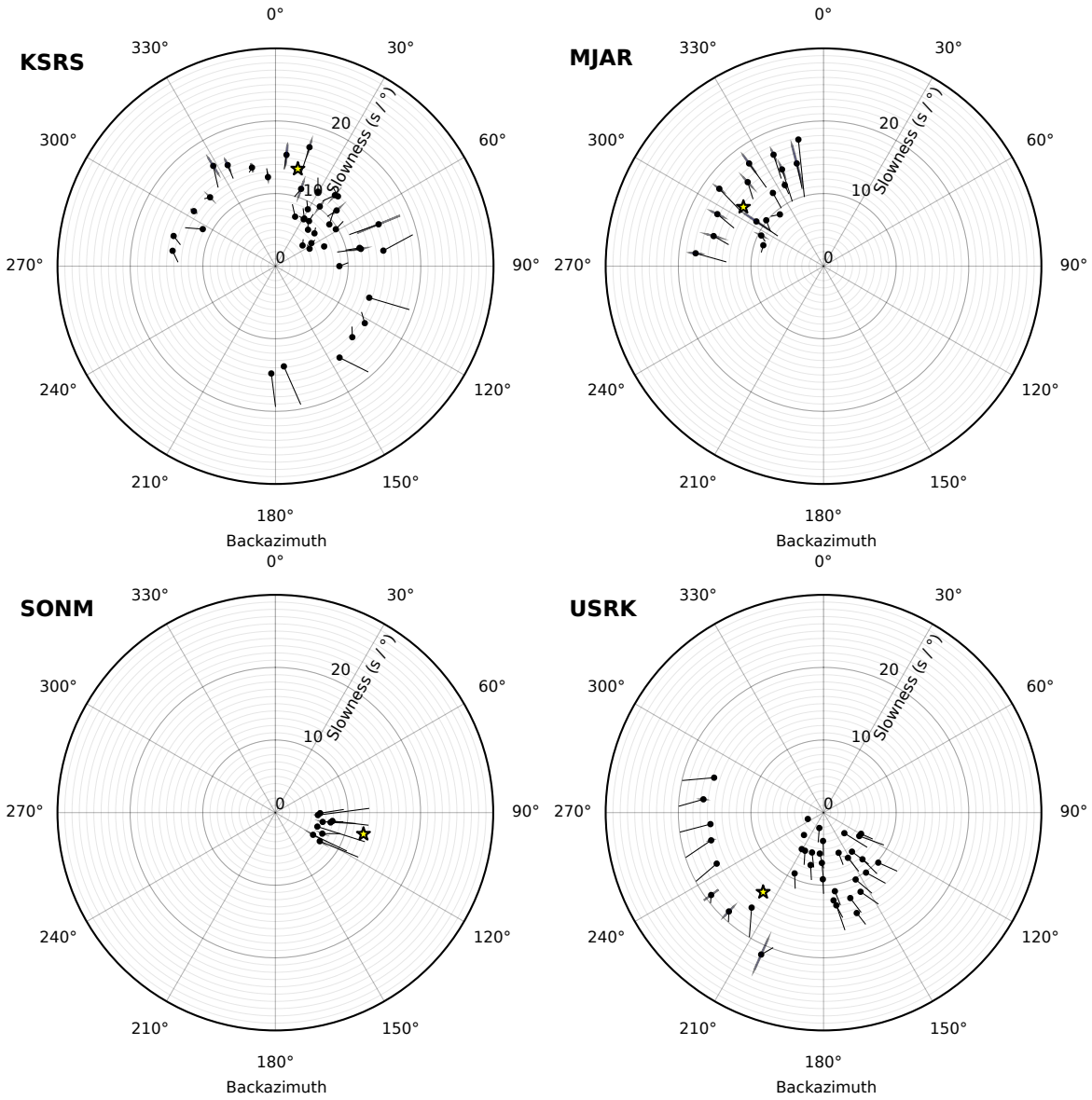

Figure 4.7: Mislocation vectors for the four arrays. Circles are plotted at the corrected slowness and backazimuth for each bin, with the vector pointing back to the centre of the bin, the theoretical slowness and backazimuth. The yellow star shows the slowness and backazimuth expected for P-wave arrivals from the Punggye-ri Nuclear Test Site. Error ellipses indicate the standard error for the SASC in each bin.

for the test site was obtained for the USRK array, so it is not shown.

For comparison, Figures 4.13– 4.20 show a selection of results for the test events. These events were chosen to represent three geographically distinct areas in the region with reasonably high seismic activity, for which SASCs were obtained for each array. Results are shown in Table **??**. Significant improvement can be seen at KSRS for Events 2 and 3, at SONM for Events 2, 3 and 4, and at USRK for Events 2 and 4. At MJAR the SASCs only yielded an improvement to the SNR

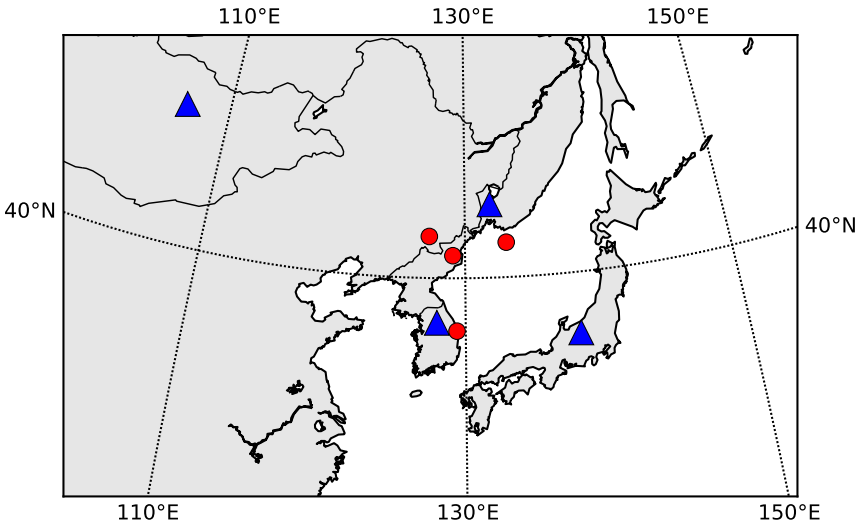

Figure 4.8: Events used in F-statistic comparison.

Table 4.4: Observed F-stat increase around the P-wave arrival for the test events after applying SASCs. Not shown where the P-wave arrival is not detected before or after corrections are made.

	KSRS	MJAR	SONM	USRK
Event 1	−58%	−56%	+8%	—
Event 2	+167%	−64%	+100%	+260%
Event 3	+3100%	—	+29%	0%
Event 4	−11%	+160%	+50%	+300%

in the stack for Event 4.

4.4.1 Application to DPRK announced nuclear tests

The SASCs for the SONM array show a good SNR improvement for the seismic signals from the four announced nuclear tests conducted 2009–2016, and no real improvement for the 2006 test (see Figure 4.22).

I use the enhanced stacks to calculate the body wave magnitude for the five events, given by

$$m_b = \log_{10} \frac{A}{T} + Q(\Delta, h) \tag{4.4}$$

where A is the maximum zero-to-peak amplitude in μm and T is the period in s as measured for the early P-wave arrival around 1 Hz. Q is the attenuation as a function of epicentral distance Δ and source depth h. For a shallow event at a distance of 17°, $Q = 5.9$ [38].

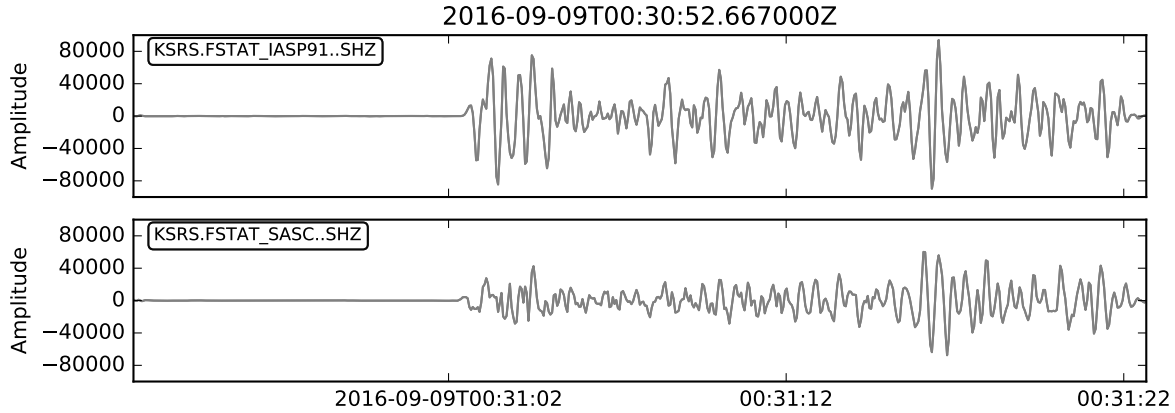

Figure 4.9: Stacks with and without SASC for the 2009 DPRK announced nuclear test at KSRS.

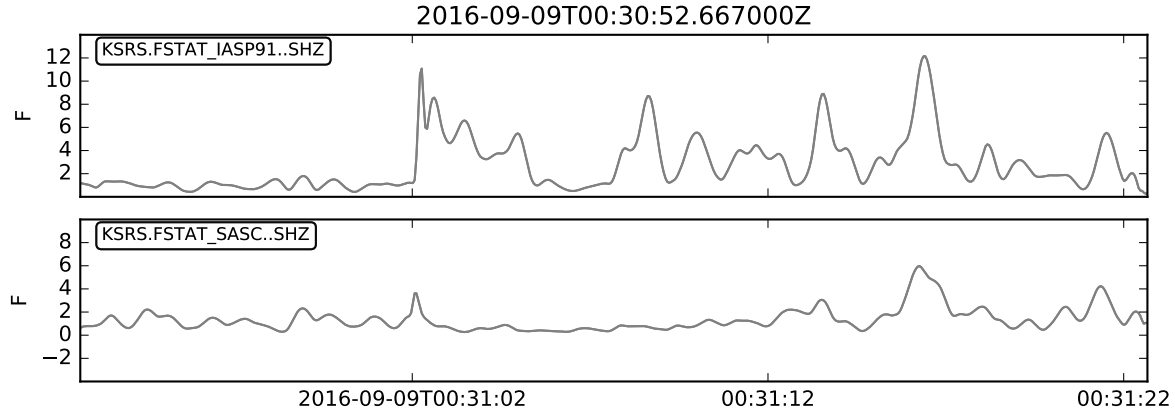

Figure 4.10: F-statistic in the stack with and without SASC for the 2009 DPRK announced nuclear test at KSRS. Despite the good SNR of the waveform, the SASC has a negative effect on the F-statistic.

To calculate m_b in the stacks, the stacked traces are filtered from 0.5 to 10 Hz, any linear trend is removed, and the instrument response is deconvolved to give the ground displacement. Example P-waves used in the magnitude calculation are pictured in Figure 4.21.

The results are given in Table **??**. The magnitudes calculated with and without SASC are compared with those reported by the IDC as well as the IDC surface wave magnitude (for the 2006 event this is from IASPEI). The last column shows a value for m_B/M_S using the SASC m_b and the IDC M_S. The magnitudes recorded at SONM are around 0.5 lower than those reported by the IDC, but those calculated from the corrected stacks are slightly larger than those using the uncorrected stack. The effect of the SASC on the m_b/M_S discriminant is not significant.

To illustrate the effect on the m_b/M_S discriminant, I plot the IDC surface wave magnitude against the body wave magnitudes I obtained from each of the stacks, and compare these with the IDC values for m_b. The results are plotted in Figure 4.23.

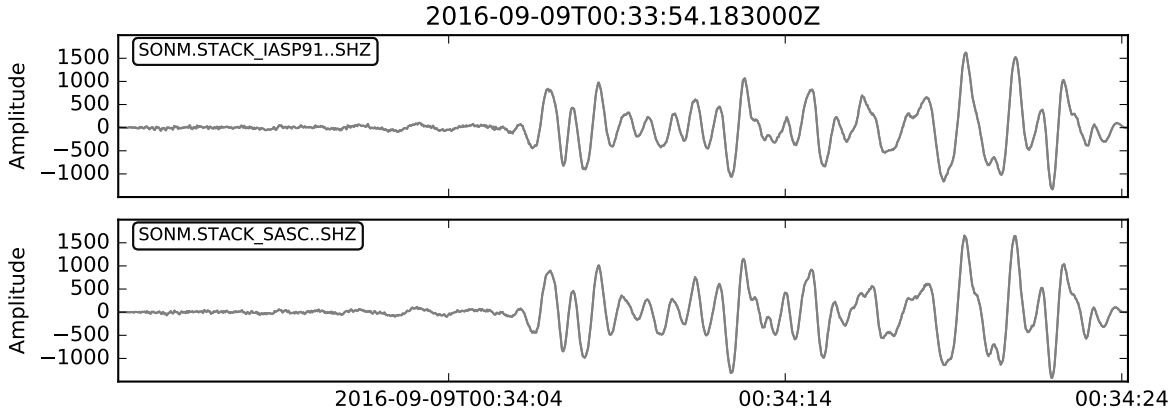

Figure 4.11: Stacks with and without SASC for the 2009 DPRK announced nuclear test at SONM.

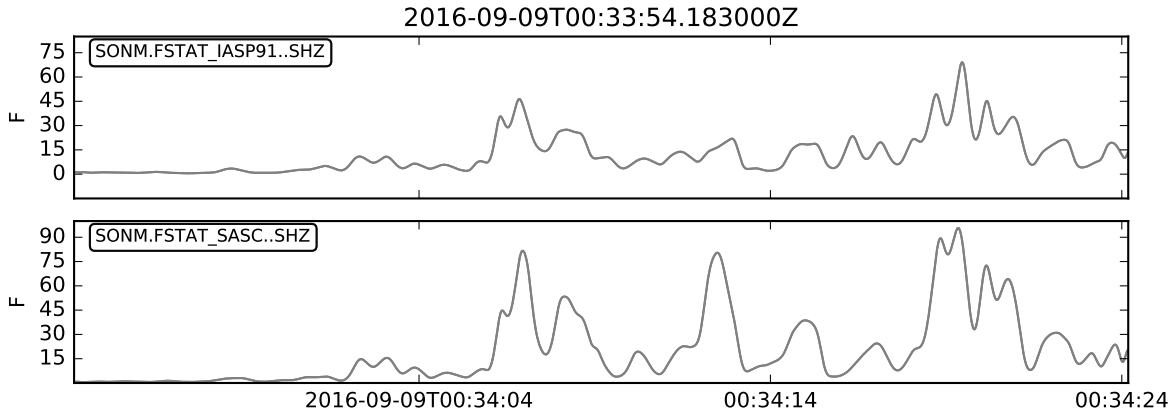

Figure 4.12: F-statistic in the stack with and without SASC for the 2009 DPRK announced nuclear test at SONM. The SASC increases the F-statistic in the stack.

For comparison, I also calculate magnitudes with and without corrections for the other arrays, for the events where the SNR is improved by the SASC. The results are shown in Table **??**. No M_S is reported by the IDC for these events, except for Event 4. The revised m_b/M_S for this event at MJAR is 1.02 ± 0.07 (without correction 1.03 ± 0.12) and at USRK 1.16 ± 0.04 (without correction 1.14 ± 0.05), well within the expected range for an earthquake.

4.5 Discussion

In this section, I will discuss the results obtained at each array, the overall pattern of the mislocation vectors and their effect on the SNR of the data. I will then look at the efficacy of the SASCs in the contexts of signal improvement, magnitude estimation and the discrimination of explosions from earthquakes. Then I will discuss the discrepancy between the magnitudes I calculate and those reported by the IDC, and possible explanations for this difference.

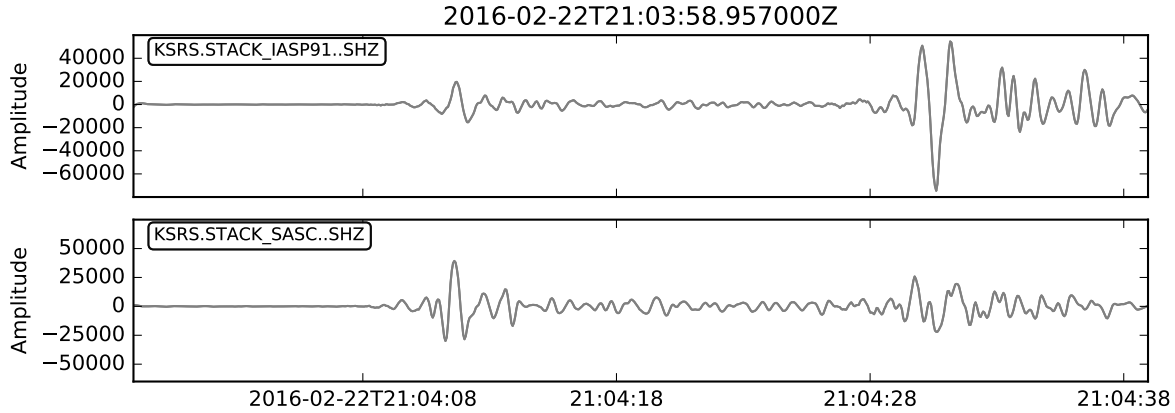

Figure 4.13: Stacks with and without SASC for the 2016 South Korea event at KSRS. The SASC visibly improves the SNR in the stack.

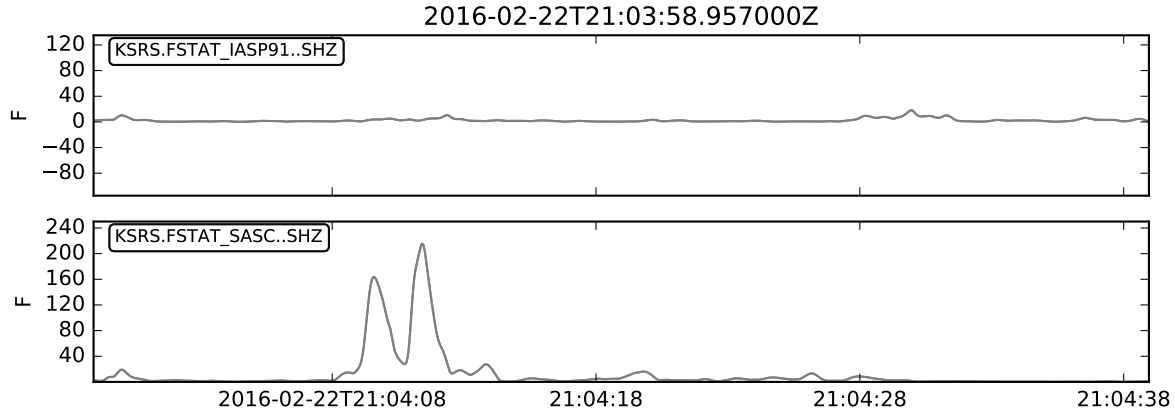

Figure 4.14: F-statistic in the stack with and without SASC for the 2016 South Korea event at KSRS. The SASC greatly increases the F-statistic in the stack.

4.5.1 SONM

The mislocation vectors for SONM are consistent across the whole region, suggesting that the velocity perturbation that is causing the apparently faster P-wave arrival lies further along the path. There is evidence for high P-wave velocity anomaly due to thicker lithosphere at around 100 km depth beneath the Gobi desert [83], a region which the ray from these events to SONM travels.

The SASCs obtained for SONM proved effective in enhancing the SNR (as measured using the F-statistic) for all the DPRK underground nuclear tests, and earthquakes in the Sea of Japan and South Korea. However the increased SNR had no significant effect on the measurement of m_b using the corrected stacks.

The SASCs obtained for SONM are consistent across the full range of backazimuths, suggesting that the cause of the shorter travel times observed for P-waves arriving from the region lies

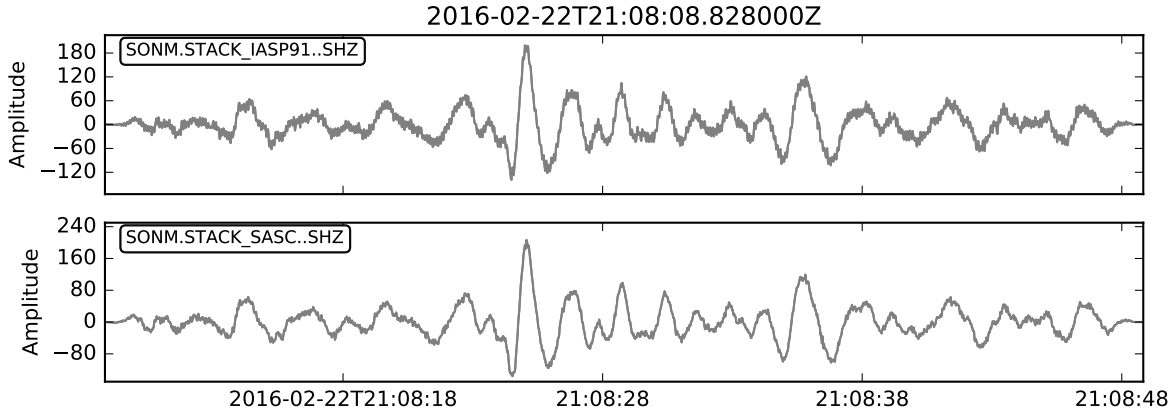

Figure 4.15: Stacks with and without SASC for the 2016 South Korea event at SONM. The SASC increases the SNR in the stack.

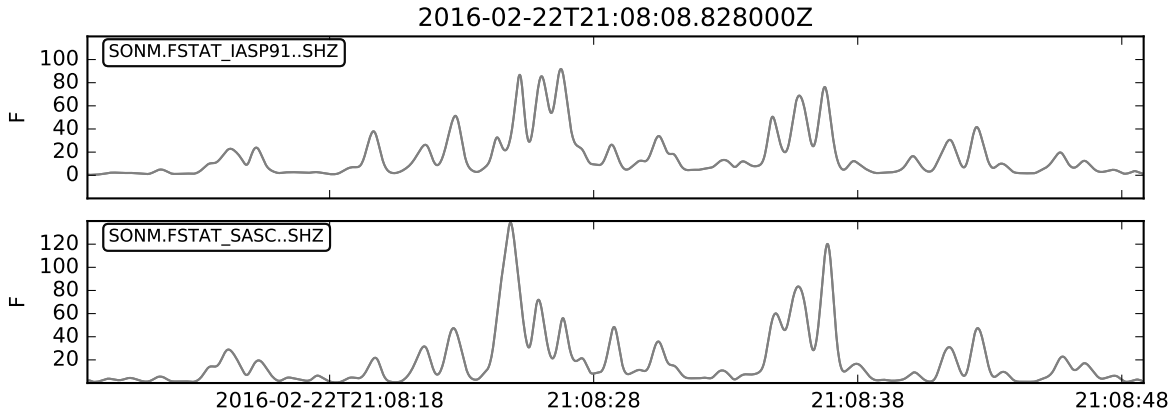

Figure 4.16: F-statistic in the stack with and without SASC for the 2016 South Korea event at SONM. The SASC slightly increases the F-statistic in the stack.

along the raypaths closer to or beneath the array site itself.

4.5.2 KSRS

At KSRS, the SASC for the DPRK Nuclear Test Site was not found to be effective in improving the SNR in the stack, although SASCs for the Sea of Japan and South Korea did increase the F-statistic measured in the stack.

The SASCs show no clear pattern based on the direction of incidence of the P-waves. As the array is sited within the event catalogue, it may be more sensitive to the complex local velocity structure of the region. The poor performance of the SASCs for the events in the DPRK Test Site and the region northeast of the array could indicate that the residuals for these events are not showing geological velocity perturbations but timing errors or other noise. However, the signal improvement seen in the Sea of Japan events to the south and west of the array may indicate a

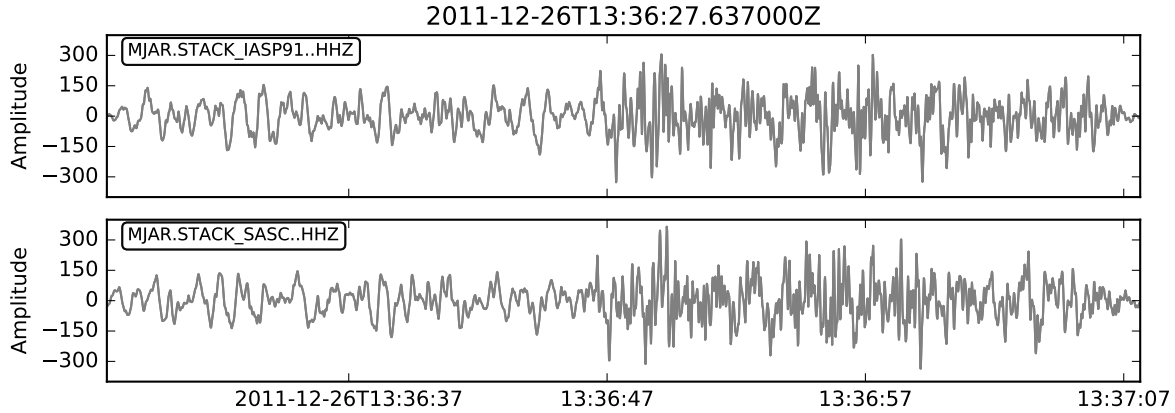

Figure 4.17: Stacks with and without SASC for the 2011 Northeastern China event at MJAR. The SASC does not improve the SNR.

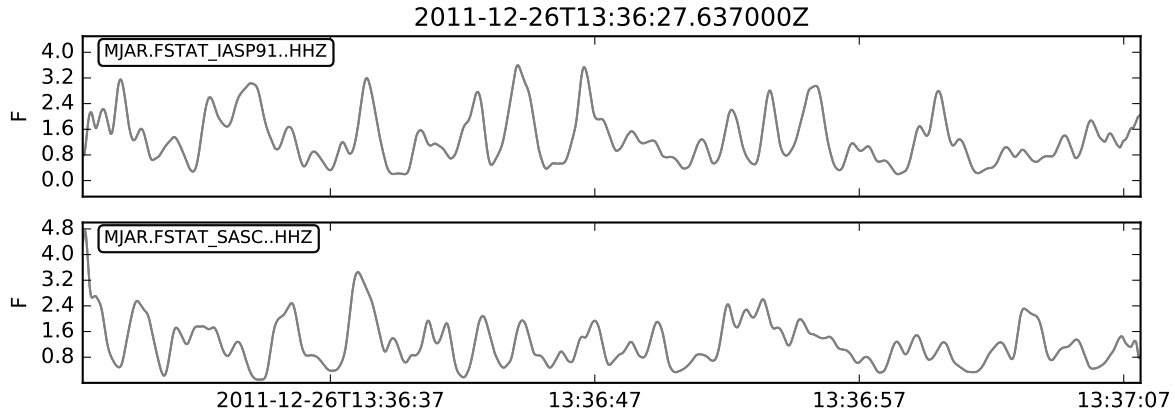

Figure 4.18: F-statistic in the stack with and without 2011 Northeastern China event at MJAR. The SASC has very little effect on the F-statistic.

region of a fast velocity perturbation along the raypath.

There is no sinusoidal pattern to the mislocation vectors that could be explained by the southwesterly dipping Moho reported by Lee *et al.* from their receiver function study in South Korea [48]. The seismic refraction survey of Cho *et al.* [17] found lower-than-average crustal velocities in the southwest of the country, which do not agree with the mislocation vectors I obtain.

4.5.3 USRK

At USRK, no SASC was obtained for the Nuclear Test Site. However, effective SASCs were obtained for other regions, including the Sea of Japan and Northeastern China.

The mislocation vectors for the northeastern China events indicate a positive velocity perturbation along the raypath, and similar faster arrivals are seen for the Sea of Japan events. This

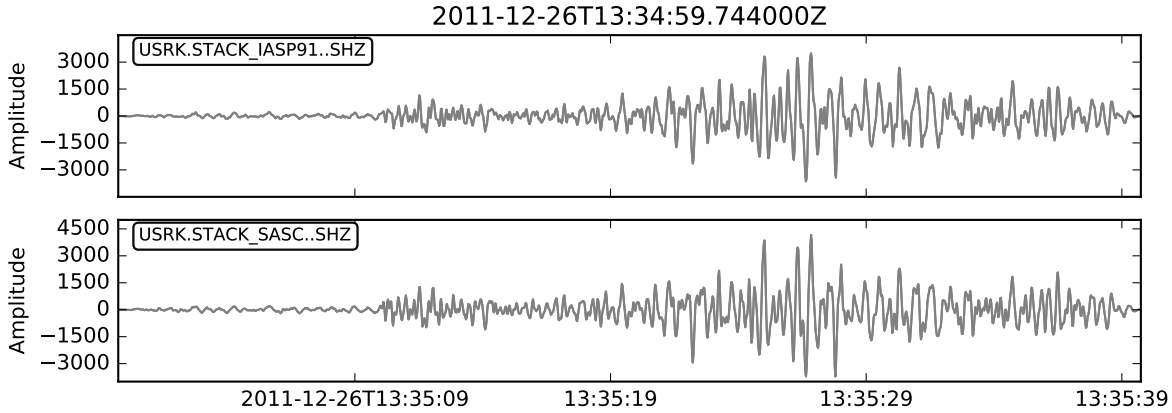

Figure 4.19: Stacks with and without SASC for the 2011 Northeastern China event at USRK.

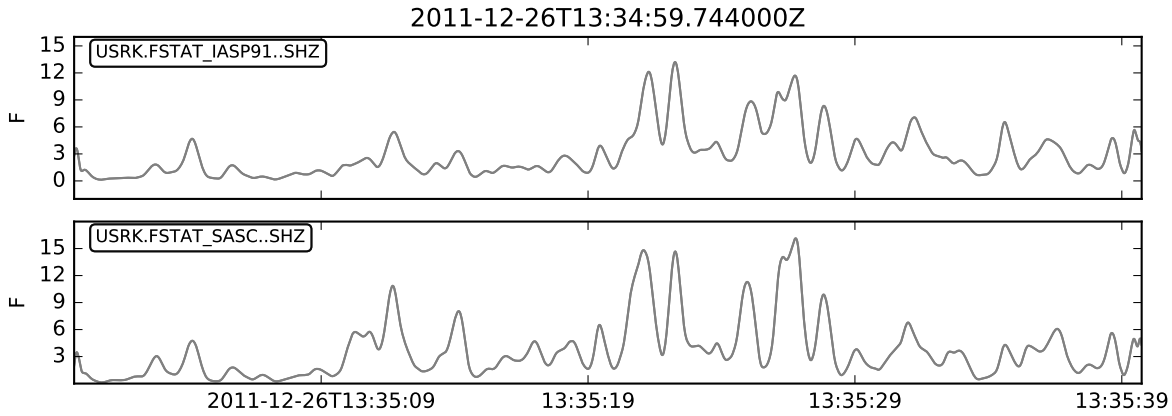

Figure 4.20: F-statistic in the stack with and without SASC for the 2011 Northeastern China event at USRK. The SASC slightly increases the F-statistic in the stack.

could be explained by a high P-wave velocity region in the upper mantle 10–60 km beneath the China/DPRK/Russia border region [61].

4.5.4 MJAR

MJAR was the only station in this study to show positive slowness corrections, i.e., slower arrivals than those predicted by *iasp91*. In their tomographic study of the region, Pei and Chen (2010) [61] found a low velocity anomaly beneath Japan.

At MJAR, none of the obtained SASCs had a positive effect on the SNR in the stack. The signal improvement of the linear stack at MJAR is limited by the incoherence of the signal waveforms between the stations of the array (Section 2.4), and it is possible that a phase-weighted stack would achieve better results with the SASCs applied.

Table 4.5: Magnitude results for the corrected and uncorrected stacks from the DPRK announced nuclear tests recorded at SONM. † for the 2006 event, the value for M_S is that recorded by IASPEI.

Event	m_b^{uncorr}	m_b^{SASC}	m_b^{IDC}	M_S
2006	3.5 ± 0.1	3.5 ± 0.1	4.0	$2.9^{\dagger}$
2009	3.9 ± 0.2	4.0 ± 0.2	4.5	3.6
2013	4.4 ± 0.4	4.4 ± 0.3	4.9	4.0
2016a	4.2 ± 0.2	4.2 ± 0.2	4.9	3.9
2016b	4.5 ± 0.2	4.5 ± 0.2	5.1	4.2

Table 4.6: Magnitude results for the corrected and uncorrected stacks for the rest events, at the arrays for which the SNR in the stack is improved by the SASCs. No estimate for M_S was available for Events 1–3, but the M_S reported for Event 4 is 3.3.

Event	m_b KSRS		m_b MJAR		m_b SONM		m_b USRK	
	iasp91	SASC	iasp91	SASC	iasp91	SASC	iasp91	SASC
2	4.6 ± 0.7	4.6 ± 0.5	—	—	4.8 ± 0.2	5.0 ± 0.3	—	—
3	5.6 ± 0.3	5.9 ± 0.4	—	—	—	—	—	—
4	—	—	3.4 ± 0.4	3.4 ± 0.2	—	—	3.8 ± 0.2	3.8 ± 0.1

4.5.5 Slowness-azimuth station corrections

Comparing my SASCs to those obtained by Bondár *et al.*, the positive corrections are not consistent with the negative corrections they obtain in the same backazimuth range. However, while they use a global teleseismic catalogue, I am sampling events closer to the array, and in a higher slowness range than those they use for their corrections. This enables me to image the velocity structure around the DPRK test site at a higher resolution. Additionally, the SONM and USRK arrays were deployed after that paper was published.

Where the SASCs have failed to improve the SNR of the stacks, this could be due to a number of factors. It is possible that the frequency band I used (0.5–5.0 Hz) is too wide, and the velocity anomalies vary significantly with frequency. In this case, different SASCs would need to be applied to different frequency ranges, and the resulting stacks filtered to a more narrow spectrum. This would lead to a reduction of information about the waveform in the stack itself.

4.5.6 Magnitude recalculation

The magnitudes I obtain for the DPRK events, measured at a single array, SONM, are an order of magnitude lower than those reported by the IDC. Magnitude estimates from a single array are limited in that they do not sample the full moment tensor from the event source. Consequently, arrays or stations in certain areas of the moment tensor will report smaller magnitudes than those in other regions—for this reason, multiple arrays and stations are used in the estimates reported in the REB. While a theoretical explosion would be expected to produce an isotropically

compressional moment tensor, those observed for the underground nuclear tests in North Korea show significant double-couple (DC) and compensated linear vector dipole (CLVD) components.

4.6 Conclusions and further work

In this chapter, I used seismic events in the region of East Asia surrounding the Punggye-ri Nuclear Test Site in North Korea to develop slowness-azimuth station corrections for four seismic arrays of the International Monitoring System (IMS). I calculated the corrections using frequency-wavenumber (fk) analysis using the VesPy software package I developed previously during my studies (see Chapter 3).

The SASCs reveal faster-than-predicted observed velocities for P-waves arriving at the SONM array in Mongolia. Making these corrections improves the signal-to-noise ratio (SNR) in the stacked waveforms for events originating from this region. However, the effect of the corrections on the measured values of the body wave magnitude m_b and thus the seismic discriminant m_b/M_S were small, and the magnitudes for the announced nuclear test events in the DPRK as measured from the stacks (before and after correction) were found to be 0.5 magnitudes smaller than those reported by the IDC.

The SASCs obtained for the KSRS and USRK arrays in South Korea and north-eastern Russia respectively were not effective for the DPRK events. However they did improve the SNR for events in the Sea of Japan and Northeastern China, suggesting perturbations to the local velocity structure. The SASCs for the MJAR array in Honshu, Japan, were not successful in improving the SNR of the waveforms.

Further development of this study could apply this method to develop SASCs for events across the entire globe, and for other IMS stations for which corrections do not currently exist. SASCs could be obtained at three-component stations for S-wave and surface wave arrivals to improve the measurement of M_S used in seismic event discrimination.

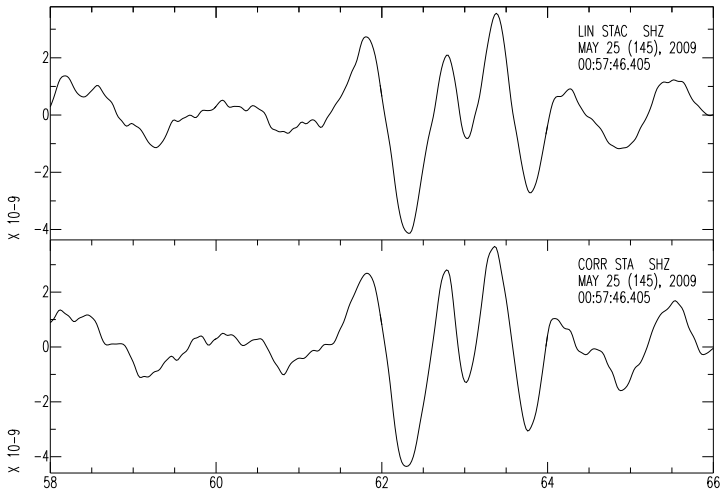

Figure 4.21: P-waves (period ~ 1 Hz) used to calculate m_B at SONM for the 2009 DPRK test event. The instrument response has been removed, linear trend and mean removed, and a Butterworth filter applied between 0.5 and 10.0 Hz.

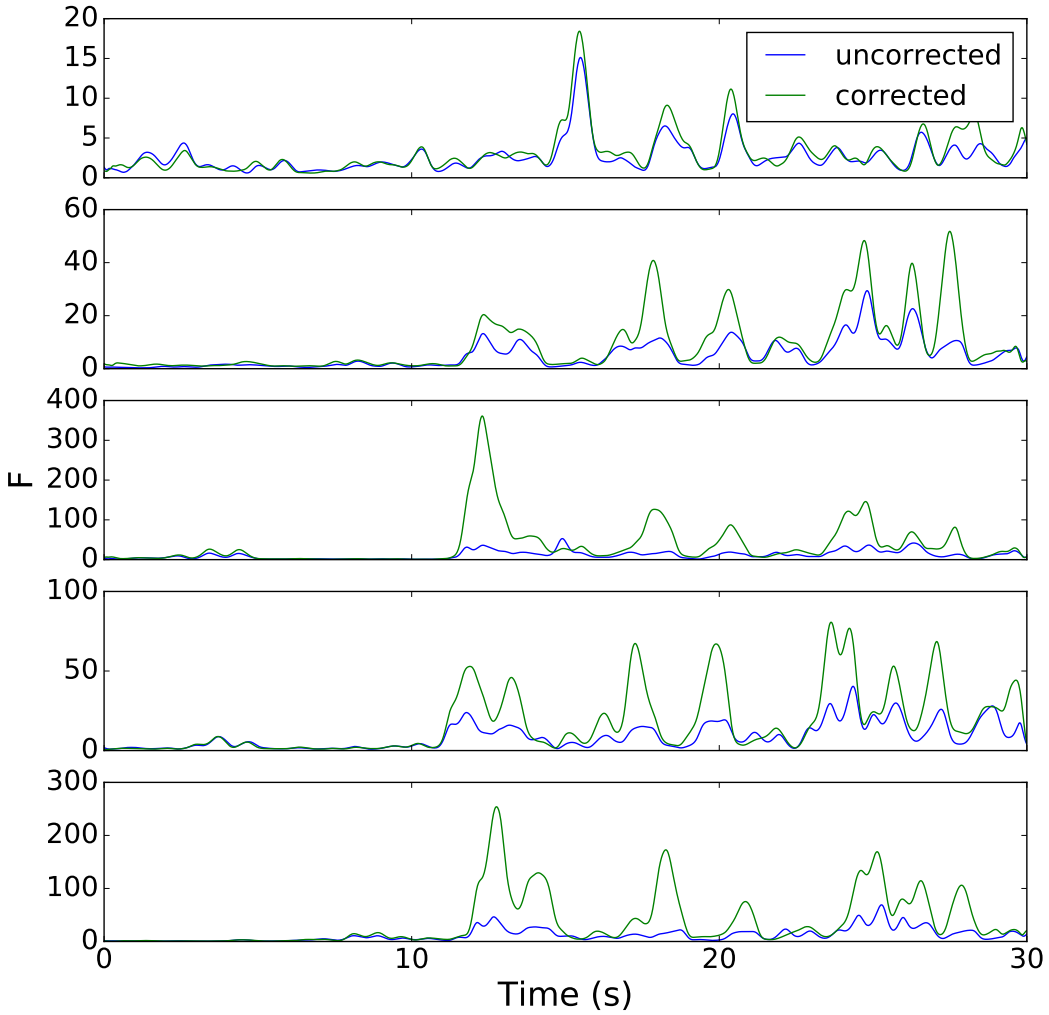

Figure 4.22: F-statistic in the stack with and without SASC recorded at SONM for the 2006–2016 announced underground nuclear tests in North Korea, showing the SNR improvement due to the SASC. In each case the P-wave expected arrival time is at 10 s

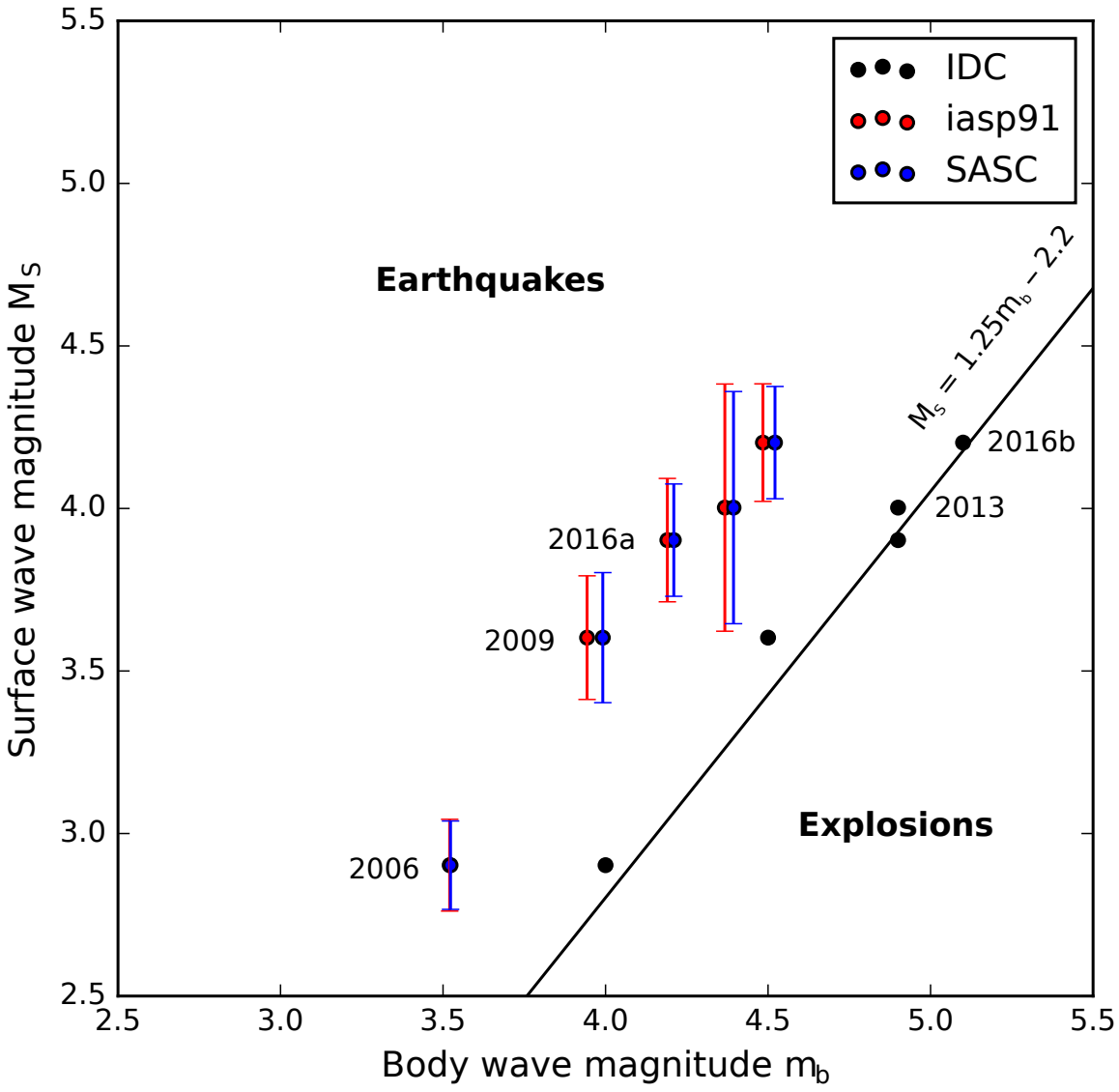

Figure 4.23: m_b vs. M_S for the five DPRK nuclear tests 2006–16 recorded at SONM. The IDC values are plotted in black, and fall on the screening line of Fisk *et al.* (2002) [30]. The m_b values for the uncorrected (iasp91) stack are plotted in red with associated error, and the SASC-corrected stack in blue. The IDC value of M_S is used in each case. The errors in the SASC are slightly reduced for all but one event, and the magnitudes closer (but still an order of magnitude less than) the IDC values. The m_b/M_S discriminant as obtained by the IDC is inconclusive for the DPRK events. The smaller magnitudes obtained using only the stacked data at SONM misidentify the events as earthquakes.

CHAPTER 5

CONCLUSIONS

In this thesis, I have described the entwined histories of seismology and nuclear testing, and how the need to monitor the development of nuclear weapons has been the driving force behind many modern processing techniques and instrumentation, benefitting the science of global seismology outside of forensic applications. I looked in detail at various array methods, including time series stacking, velocity spectral analysis and frequency-wavenumber analysis and their applications in improving the quality of seismic signals recorded at arrays, and in detecting, locating and identifying events.

The central product of my PhD project is the VesPy software package, which I have made freely available to seismologists to use in their own array studies. The development of this package was motivated by how useful I and my colleagues have found the Python programming language in seismic studies, and to add to a library of existing scientific and seismological Python code, such as SciPy and ObsPy, with a package of dedicated functions for seismic array analysis. A full description of the package, including example applications, is in Chapter 3 and documentation for installing VesPy can be found in Appendix A. The examples in this thesis beyond forensic seismology include detecting core phases using the `vespagram` function and calculating slowness anomalies using the `fk_analysis` function.

VesPy was used in the data processing for the study of stacking techniques in Chapter 2 and the regional study of the area around the North Korean test site in Chapter 4. In the latter chapter, I focused on a particular application useful for forensic seismology, developing slowness-azimuth corrections for arrays of the International Monitoring System (IMS) in order to improve the monitoring capability of those arrays for any future nuclear testing in the region. These corrections were then applied to nuclear test events in North Korea, and were found to increase the signal-to-noise ratio in the stacked waveform, changing the body-wave magnitude

estimate, and reducing the uncertainty in this calculation.

In this chapter, I propose a workflow for the study of future announced or clandestine nuclear testing anywhere in the world, following the results of my work described in this thesis. Finally, I make suggestions for how this work might be extended beyond the scope of this project.

5.1 Recommended workflow for nuclear test monitoring at seismic arrays

Seismic monitoring of potential underground nuclear tests remains a crucial part of the International Monitoring System. Seismic arrays allow the signal-to-noise ratio of data from these events to be improved towards two main purposes useful to a forensic seismologist: a) signal detection, and b) studies of the seismic waveform, including discrimination of the type of event.

5.1.1 Which seismic arrays to use?

For signal detection, there are two factors to consider when choosing an array: the potential signal improvement, and the slowness and azimuth resolution. Both are desirable to detect small magnitude events and to be able to determine their origin precisely. For the former, a large number of stations in the array is useful, but to ensure the latter, the array geometry must be considered. To ensure a good coverage of backazimuths, a circular array (such as TORD in Niger [29]) or spiral-armed array (PSA, Australia) is ideal. Slowness resolution depends on the aperture of the array, with a larger aperture giving a higher resolution.

For detection at teleseismic distances, an array with a 10–20 km circular or spiral aperture yields, such as WRA or PSA give good results. Coherent arrays of this geometry also give good results for waveform studies.

5.1.2 Which stacking technique to use?

As described in the Conclusions to Chapter 2, the choice of stacking technique depends on whether the shape of the waveform needs to be preserved or not. It also depends on the properties of the array itself, with very large arrays being unsuitable for stacking where the plane-wave approximation does not apply.

Before deciding on a stacking technique it is useful to know the coherence properties of the array, i.e. how similar the signal is between the different array elements. This can be quantified using cross-correlation, or by calculating the semblance in the linear stack (e.g., using the `semblance` function in VesPy.)

- If detection is paramount, the eighth-order nth-root stack gives the best SNR enhancement and slowness resolution in velocity spectral analysis.

- For studies where the shape of the waveform is important, such as magnitude or amplitude calculations, receiver functions, or correlation studies, the first-order phase-weighted stack performs well even at incoherent arrays. Further SNR improvement is possible at coherent arrays using the adaptive stacking technique of Rawlinson *et al.* [64].

5.1.3 Best practice for developing slowness-azimuth station corrections

Slowness-azimuth station corrections (SASCs) improve the signal enhancement of the linear stack (and other techniques) by accounting for velocity perturbations under the array, and along the raypath from the source to the receiver. SASCs can be calculated for a particular array by determining the slowness and azimuth at which seismic phases are observed to arrive, and how this differs from those predicted by the *a priori* velocity model. These residuals must be calculated for a large catalogue of events, which can be regional or global in scale depending on the area(s) of interest for the corrections, and their distance to the array.

The seismic events in the catalogue used to develop the SASCs should resemble as closely as possible the underground explosions one is aiming to detect i.e., they will be predominantly shallow events (~ 10 km depth) in the magnitude range 3.0–5.0. Larger nuclear explosions, such as the magnitude 6.1 seismic event following the 2017 DPRK nuclear test, generate signals with higher SNR, and so are easier to detect without stacking. For regions with established test sites, such as Punggye-ri in the DPRK, or the Nevada Test Site, there is a historical catalogue of underground explosion events that can be used to train the SASCs. However, in an area where no previous tests have been detected, the event catalogue will be constrained by the region's seismicity. A region of high shallow seismicity will be easier to produce an SASC for, but any clandestine nuclear tests conducted there will need to be accurately discriminated from natural earthquakes.

In order to calculate the residuals, a combination of fk analysis and the F-statistic (both implemented in VesPy) is effective at detecting the arrival of a seismic phase with a high slowness resolution. Before fk analysis, the seismic waveforms need to be trimmed to a narrow time window of a few seconds around the predicted phase arrival, and filtered to the required frequency range. SASCs calculated for narrow frequency bands (instead of the broader 0.5–5.0 Hz band used in this study) may yield better results for arrays very near (<5°) from the study region.

Following the method of Bondár *et al.* [10], the events are binned by backazimuth and slowness, and the median residual within each bin taken as the SASC. The size of the bin depends on the scale of the study and the distance from the source to the array. At even far-regional distances (~ 20°), the SASCs can be similar across a region, as for example observed at SONM for the region around the Korean peninsula, at a distance of 17° (see Figure 4.7). The observed pattern of mislocation vectors can indicate the source of the anomaly: consistent corrections for a particular range of backazimuths and slowness might indicate a velocity perturbation along a particular raypath, whereas a systematic shift across the whole array might be due to

miscalibration of the array's instrumentation. Sinusoidal slowness perturbations are known to be due to a dipping Moho beneath the array [31].

5.2 Future work

While this study has focused on forensic seismology as an application for array methods, many of the techniques can be applied to global seismology. Slowness-azimuth station corrections can be obtained not just for P-waves but for S-waves and surface waves at regional and teleseismic distances to more fully describe the velocity structure under the array. As well as slowness corrections, static corrections to the arrival time and amplitude for each station in the array can be calculated using the adaptive stacking technique.

Another common application of array seismology is in exploration geophysics, and microseismic event detection in reservoirs. The various stacking functions in VesPy can be applied to microseismic data to compare how the optimal stacking technique differs for very small seismic events occurring at short distances from the receiver. Power statistics like the semblance and F-statistic can play a useful role in stacking and signal detection in microseismic studies.

The functions in the VesPy package can be applied to seismic array analysis for many purposes, and the vespagram and fk analysis functions are useful for detecting and relocating events for tectonic studies.

APPENDIX A

VesPy Documentation

A.1 Download and installation

The VesPy software package is primarily available through GitHub at http://github.com/NeilWilkins/VesPy. It is freely distributed under the MIT licence and the full source code can be downloaded (cloned) to a local repository.

Installation can be done via Python using the included `setup.py` installation script. Run the following command within the `vespy` directory:-

```
python setup.py install
```

A.2 Using the code

Individual functions or complete modules can be imported for use in Python (2 or 3), e.g.

```
from vespy.fk import fk_analysis
```

Detailed docstrings are provided for each function, describing input parameters and return values.

BIBLIOGRAPHY

[1] R. D. ADAMS, *Reflections from discontinuities beneath Antarctica*, Bulletin of the Seismological Society of America, 61 (1971), pp. 1441–1451.

[2] M. AUER AND M. K. PRIOR, *A new era of nuclear test verification*, Physics Today, 67 (2014), pp. 39–44.

[3] R. G. BARON, J. H. COCHRANE, J. R. FISCHER, R. L. FISHER, S. G. FRANCISCO, S. D. LORENZ, J. R. MADIGAN, F. ROSENTHAL, K. C. SIH, W. D. SMITH, C. P. SNIVELY, N. TINANOFF, E. G. UNGAR, W. VANDERKULK, T. P. WALKER, AND J. S. WARDEN, *LASA signal processing, simulation and communications study*, tech. rep., Directorate of Planning and Technology, Electronic Systems Division, Air Force Systems Command, United States Air Force, 1967.

[4] M. BEYREUTHER, R. BARSCH, L. KRISCHER, T. MEGIES, AND Y. BEHR, *ObsPy : A Python Toolbox for Seismology*, Seismological Research Letters, 81 (2010), pp. 530–533.

[5] J. W. BIRTILL AND F. E. WHITEWAY, *The Application of Phased Arrays to the Analysis of Seismic Body Waves*, Philosophical Transactions of the Royal Society of London, 258 (1965), pp. 421–493.

[6] S. BLAKESLEE, *In Remotest Nevada, a Joint U.S. and Soviet Test*, aug 1988.

[7] R. BLANDFORD, *An Automatic Event Detector at the Tonto Forest Seismic Observatory*, Geophysics, 39 (1974), pp. 633–643.

[8] B. A. BOLT, *Nuclear Explosions and Earthquakes: The Parted Veil*, W. H. Freeman and Co., 1976.

[9] B. A. BOLT, *The detection of PKIIKP and damping in the inner core*, Annals of Geophysics, 30 (1977), pp. 507–520.

[10] I. BONDÁR, R. G. NORTH, AND G. BELL, *Teleseismic Slowness-Azimuth Station Corrections for the International Monitoring System Seismic Network*, Bulletin of the Seismological Society of America, 89 (1999), pp. 989–1003.

[11] D. BOWERS, *Using the F-Detector to help interpret P-seismograms recorded by seismometer arrays*, in Proceedings of the Seismic Research Symposium, New Orleans, USA, 2000, p. 10.

[12] D. BOWERS AND N. D. SELBY, *Forensic Seismology and the Comprehensive Nuclear-Test-Ban Treaty*, Annual Review of Earth and Planetary Sciences, 37 (2009), pp. 209–236.

[13] W. BURR, *Documents on the US Atomic Energy Detection System*, 2017.

[14] J. CAPON, *High-Resolution Frequency-Wavenumber Spectrum Analysis*, Proceedings of the IEEE, 57 (1969), pp. 1408–1418.

[15] J. CAPON, R. J. GREENFIELD, AND R. T. LACOSS, *Off-line signal processing results for the Large Aperture Seismic Array*, tech. rep., Lincoln Laboratory, Massachussets Institute of Technology, 1966.

[16] E. J. CHAVES, T. LAY, AND D. P. VOYTAN, *Yield Estimate (230 kt) for a Mueller-Murphy Model of the 3 September 2017, North Korean Nuclear Test (mbNEIC = 6.3) From Teleseismic Broadband P Waves Assuming Extensive Near-Source Damage*, Geophysical Research Letters, (2018).

[17] H. M. CHO, C. E. BAAG, J. M. LEE, W. M. MOON, H. JUNG, K. Y. KIM, AND I. ASUDEH, *Crustal velocity structure across the southern Korean Peninsula from seismic refraction survey*, Geophysical Research Letters, 33 (2006), pp. 6–9.

[18] J. COYNE, D. BOBROV, P. BORMANN, E. DURAN, P. GRENARD, G. HARALABUS, I. KITOV, AND Y. STAROVOIT, *CTBTO: Goals, networks, data analysis and data availability*, in New Manual of Seismological Observatory Practice 2, vol. 1, IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 2012, ch. 15, pp. 1–34.

[19] J. A. CROCKER, *Seismic waves from A-bombs detonated over a land mass*, tech. rep., U.S. Air Force Office for Atomic Energy, 1952.

[20] H. P. CROTWELL, T. J. OWENS, AND J. RITSEMA, *The TauP Toolkit: Flexible Seismic Travel-time and Ray-path Utilities*, Seismological Research Letters, 70 (1999), pp. 154–160.

[21] CTBTO, *CTBTO hydroacoustic data used to aid search for missing submarine ARA San Juan*, 2017.

[22] D. DAVIES, E. J. KELLY, AND J. R. FILSON, *Vespa Process for Analysis of Seismic Signals*, Nature Physical Science, 232 (1971), pp. 8–13.

[23] L. E. DE GEER, *Radionuclide Evidence for Low-Yield Nuclear Testing in North Korea in April/May 2010*, Science and Global Security, 20 (2012), pp. 1–29.

[24] ——, *The Great 2010 DPRK Nuclear Test Debate: Summarizing the Evidence of a Low-Yield Nuclear Test Carried out in North Korea in May 2010*, 2015.

[25] J. DENOYER, *Depth of focus discrimination by crustal phases for NTS nuclear explosions*, tech. rep., Institute for Defense Analyses, Research and Engineering Support Division, 1963.

[26] W. H. DIMENT, S. W. STEWART, AND J. C. ROLLER, *Crustal structure from the Nevada test site to Kingman, Arizona, from seismic and gravity observations*, Journal of Geophysical Research, 66 (1961), pp. 201–214.

[27] A. DOUGLAS, *Forensic Seismology and Nuclear Test Bans*, Cambridge University Press, 2013.

[28] M. ELGABRY, I. KORRAT, H. HUSSEIN, AND I. HAMAMA, *Infrasound detection of meteors*, NRIAG Journal of Astronomy and Geophysics, 6 (2017), pp. 68–80.

[29] C. ESTABROOK, B. BERGSSON, S. SOUMANA, O. BOUREIMA, AND M. MOUMOUNI, *Results from IMS Seismic Array in Niger*, in Geophysical Research Abstracts, EGU General Assembly 2009, vol. 11, 2009.

[30] M. D. FISK, D. JEPSEN, AND J. R. MURPHY, *Experimental seismic event-screening criteria at the prototype international data center*, Pure and Applied Geophysics, 159 (2002), pp. 865–888.

[31] M. FLANAGAN, S. MYERS, AND N. SIMMONS, *Model-based corrections to observed back azimuth and slowness observations from a dipping Mohorovicic discontinuity*, Monitoring Research Review, Albuquerque, NM, United States, Sep 18 - Sep 20, 2012, (2012), pp. LLNL–CONF–563592.

[32] S. GAN, S. WANG, Y. CHEN, S. QU, AND S. ZU, *Velocity analysis of simultaneous-source data using high-resolution semblance-Coping with the strong noise*, Geophysical Journal International, 204 (2016), pp. 768–779.

[33] S. GIBBONS, F. PABIAN, S. NÄSHOLM, T. KVÆRNA, AND S. MYKKELTVEIT, *Accurate relative location estimates for the North Korean nuclear tests using empirical slowness corrections*, Geophysical Journal International, 208 (2017), pp. 101–117.

[34] S. J. GIBBONS AND F. RINGDAL, *Seismic monitoring of the North Korea nuclear test site using a multichannel correlation detector*, IEEE Transactions on Geoscience and Remote Sensing, 50 (2012), pp. 1897–1909.

[35] S. J. GIBBONS, F. RINGDAL, AND T. KVÆRNA, *Detection and characterization of seismic phases using continuous spectral estimation on incoherent and partially coherent arrays*, Geophysical Journal International, 172 (2008), pp. 405–421.

[36] P. E. GREEN JR. AND R. V. WOOD JR., *Large Aperture Seismic Array Capabilities*, Tech. Rep. 902, Lincoln Laboratory, Massachusetts Institute of Technology, 1966.

[37] B. GUTENBERG, *Interpretation of Records Obtained from the New Mexico Atomic Bomb Test*, Bulletin of the Seismological Society of America, 36 (1945), pp. 327–330.

[38] B. GUTENBERG AND C. F. RICHTER, *Magnitude and energy of earthquakes*, Annals of Geophysics, 9 (1956), pp. 1–15.

[39] D. V. HELMBERGER AND D. G. HARKRIDER, *Seismic Source Descriptions of Underground Explosions and a Depth Discriminate*, Geophysical Journal of the Royal Astronomical Society, 31 (1972), pp. 45–66.

[40] J. D. HUNTER AND B. J. D. HUNTER, *Matplotlib : A 2D Graphics Environment*, Scientific Programming, 90 (2011), pp. 90–95.

[41] ISC, *Reference Event Bulletin*.

[42] E. JACOBEIT, C. THOMAS, AND F. VERNON, *Influence of station topography and Moho depth on the mislocation vectors for the Kyrgyz Broadband Seismic Network (KNET)*, Geophysical Journal International, 193 (2013), pp. 949–959.

[43] E. KANASEWICH, C. HEMMINGS, AND T. ALPASLAN, *Nth-root stack nonlinear multichannel filter*, Geophysics, 38 (1973), pp. 327–338.

[44] B. L. N. KENNETT, *Stacking three-component seismograms*, Geophysical Journal International, 141 (2000), pp. 263–269.

[45] B. L. N. KENNETT AND E. R. ENGDAHL, *Travel times for global earthquake location and phase identification*, Geophysical Journal International, 105 (1991), pp. 429–465.

[46] B. L. N. KENNETT, E. R. ENGDAHL, AND R. BULAND, *Constraints on seismic velocities in the Earth from travel times*, Geophysical Journal International, 122 (1995), pp. 108–124.

[47] B. L. N. KENNETT, J. STIPČEVIĆ, AND A. GORBATOV, *Spiral-Arm Seismic Arrays*, Bulletin of the Seismological Society of America, 105 (2015), pp. 2109–2116.

[48] D. H. LEE, J. M. LEE, H. M. CHO, AND T. S. KANG, *3D crustal velocity structure beneath the broadband seismic array in the Gyeongju area of Korea by receiver function analyses*, Tectonophysics, 689 (2016), pp. 89–106.

[49] J. LEWIS, *The Great 2010 DPRK Nuclear Test Debate*, 2015.

[50] R. C. LIEBERMANN AND P. W. POMEROY, *Relative excitation of surface waves by earthquakes and underground explosions*, Journal of Geophysical Research, 74 (1969), pp. 1575–1590.

[51] E. LIU, L. ZHU, A. GOVINDA RAJ, J. H. MCCLELLAN, A. AL-SHUHAIL, S. L. I. KAKA, AND N. IQBAL, *Microseismic events enhancement and detection in sensor arrays using autocorrelation-based filtering*, Geophysical Prospecting, 65 (2017), pp. 1496–1509.

[52] L. LIU, S. WU, J. CAO, F. XIE, Q. SHI, C. ZHANG, H. TANG, X. HE, R. ZHANG, L. CHEN, G. WEI, Z. ZHANG, J. ZHANG, AND H. DANG, *Monitoring of atmospheric radionuclides from the Fukushima nuclear accident and assessing their impact on Xi'an, China*, Chinese Science Bulletin, 58 (2013), pp. 1585–1591.

[53] J. L. LIVERMAN, *Final Environmental Impact Statement: Nevada Test Site*, tech. rep., Energy Research & Development Administration, 1977.

[54] N. S. NEIDELL AND M. T. TANER, *Semblance and other coherency measures for multichannel data*, Geophysics, 36 (1971), pp. 482–497.

[55] NORSAR, *NORSAR becomes the first non-US node on ARPANET, the predecessor to today's Internet*, 2017.

[56] ——, *The nuclear explosion in North Korea on 3 September 2017: A revised magnitude assessment*, 2017.

[57] NTI, *The CNS North Korea Missile Test Database*, 2018.

[58] N. ORESKES, *Plate Tectonics: An Insider's History Of The Modern Theory Of The Earth*, CRC Press, 2018.

[59] F. PABIAN, *Destruction at North Korea's Nuclear Test Site: A Review in Photos*, 2018.

[60] M. E. PASYANOS AND W. R. WALTER, *Improvements to regional explosion identification using attenuation models of the lithosphere*, Geophysical Research Letters, 36 (2009), pp. 1–5.

[61] S. PEI AND Y. J. CHEN, *Tomographic structure of East Asia : I . No fast (slab) anomalies beneath 660 km discontinuity*, Earthquake Science, 23 (2010), pp. 597–611.

[62] B. J. PETERSON, C. R. HUTT, S. JEWELL, AND U. S. G. SURVEY, *World-Wide Standardized Seismograph Network : A Data Users Guide*, tech. rep., United States Geological Survey, 2014.

[63] T. PRASTOWO AND MADLAZIM, *Real-time detection and characterization of nuclear explosion using broadband analyses of regional seismic stations*, Journal of Physics: Conference Series, 953 (2018).

[64] N. RAWLINSON AND B. L. N. KENNETT, *Rapid estimation of relative and absolute delay times across a network by adaptive stacking*, Geophysical Journal International, 157 (2004), pp. 332–340.

[65] E. A. ROBINSON, *Seismic arrays for the detection of nuclear explosions*, tech. rep., Massachusetts Institute of Technology, 1964.

[66] S. ROST AND C. THOMAS, *Array seismology: Methods and applications*, Reviews of Geophysics, 40 (2002).

[67] E. RUIGROK, S. J. GIBBONS, AND K. WAPENAAR, *Cross-correlation beamforming*, Journal of Seismology, 21 (2017), pp. 495–508.

[68] T. RYBERG AND M. WEBER, *Receiver function arrays : a reflection seismic approach*, Geophysical Journal International, 141 (2000), pp. 1–11.

[69] M. SCHIMMEL AND H. PAULSSEN, *Noise reduction and detection of weak, coherent signals through phase weighted stacks*, Geophysical Journal International, 130 (1997), pp. 497–505.

[70] J. SCHWEITZER, J. FYEN, S. MYKKELTVEIT, S. J. GIBBONS, M. PIRLI, D. KÜHN, AND T. KVÆRNA, *Seismic Arrays*, in New Manual of Seismological Observatory Practice 2, no. December, IASPEI, GFZ German Research Centre for Geosciences, Potsdam, 2011, ch. 9, pp. 1–70.

[71] J. SCHWEITZER, J. FYEN, S. MYKKELTVEIT, AND T. KVÆRNA, *Seismic Arrays*, in New Manual of Seismological Observatory Practice 2, 2012, ch. 9, pp. 1–80.

[72] N. D. SELBY, *Application of a generalized F detector at a seismometer array*, Bulletin of the Seismological Society of America, 98 (2008), pp. 2469–2481.

[73] R. E. SHERIFF AND L. P. GELDART, *Exploration Seismology*, Cambridge University Press, 1995.

[74] P. G. SILVER AND W. W. CHAN, *Shear wave splitting and subcontinental mantle deformation*, Journal of Geophysical Research: Solid Earth, 96 (1991), pp. 16429–16454.

[75] F. STANĚK, D. ANIKIEV, J. VALENTA, AND L. EISNER, *Semblance for microseismic event detection*, Geophysical Journal International, 201 (2015), pp. 1362–1369.

[76] J. L. STEVENS AND S. M. DAY, *The physical basis of mb : Ms and variable frequency magnitude methods for earthquake/explosion discrimination*, Journal of Geophysical Research: Solid Earth, 90 (1985), pp. 3009–3020.

[77] P. STUBBS, *Simpler detection of underground bomb tests*, New Scientist, (1962).

[78] L. R. SYKES, *Dealing with Decoupled Nuclear Explosions under a Comprehensive Test Ban Treaty*, in Monitoring a Comprehensive Test Ban Treaty, E. S. Husebye and A. M. Dainty, eds., Springer Netherlands, Dordrecht, 1996, pp. 247–293.

[79] S. R. TAYLOR, *Analysis of High-Frequency Pg / Lg Ratios from NTS Explosions*, Bulletin of the Seismological Society of America, 86 (1996), pp. 1042–1053.

[80] G. E. VAN DER VINK, *The containment of underground nuclear explosions*, tech. rep., Congress of the U.S., Office of Technology Assessment, 1989.

[81] S. VAN DER WALT, S. C. COLBERT, AND G. VAROQUAUX, *The NumPy Array: A Struture for Efficient Numerical Computation*, Computing in Science & Engineering, 13 (2011), pp. 22–30.

[82] F. E. WHITEWAY, *The Use of Arrays for Earthquake Seismology*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 290 (1966), pp. 328–342.

[83] F. ZHANG, Q. WU, S. P. GRAND, Y. LI, M. GAO, S. DEMBEREL, M. ULZIIBAT, AND U. SUKHBAATAR, *Seismic velocity variations beneath central Mongolia: Evidence for upper mantle plumes?*, Earth and Planetary Science Letters, 459 (2017), pp. 406–416.

[84] M. ZHANG AND L. WEN, *Seismological Evidence for a Low-Yield Nuclear Test on 12 May 2010 in North Korea*, Seismological Research Letters, 86 (2015), pp. 138–145.

[85] L.-F. ZHAO, X.-B. XIE, W.-M. WANG, J.-L. HAO, AND Z.-X. YAO, *Seismological investigation of the 2016 January 6 North Korean underground nuclear test*, Geophysical Journal International, 206 (2016), pp. 1487–1491.